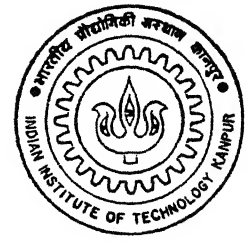
LOCAL FEATURE EXTRACTION USING WAVELET PACKETS

by

Major Sharad Shukla

EE
1996
M
SHU
LOC

TH
EE|1996|M
Sh 92 l



DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

FEBRUARY, 1996

LOCAL FEATURE EXTRACTION USING WAVELET PACKETS

A Thesis Submitted

in Partial Fulfilment of the Requirements

for the Degree of

MASTER OF TECHNOLOGY

by

MAJOR SHARAD SHUKLA

to the

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

February 1996

21 MAR 1996
CENTRAL LIBRARY
I. I. T., KANPUR
Acc. No. A. 121214

EE-1996-M-SHU-LOC

Certificate

It is certified that the work contained in the thesis entitled **LOCAL FEATURE EXTRACTION USING WAVELET PACKETS**, by Major Sharad Shukla, has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

Dr. Sumana Gupta
Professor
Department of Electrical Engineering
I.I.T. Kanpur

Dedicated to my family

Acknowledgement

I would take this opportunity to express my sincere gratitude towards Dr Sumana Gupta, my thesis supervisor for her invaluable guidance. I would also like to thank Maj B V B Prasad, Sanjay G. Joshi and Gomathi Sankar, who were a great support during my thesis .

Finally, I would like to express my gratitude towards the Indian Army and Corps of Signals in particular for having given me this opportunity.

Major Sharad Shukla

Abstract

Feature extraction forms a very important part of image processing. In this thesis we have considered various aspects of feature extraction, viz, compression, noise removal, edge charecterisation and local rescaling. Emphasis here has been on image compression and noise removal. We have used wavelet packets as the tool for achieving the same.

For the purpose of image compression, we have used 'best-basis' algorithm using entropy as the cost function. Selection of the 'best-basis' is followed by a thresholding based on an energy criterion. A variety of mother wavelet functions have been used and it has been observed that the compression ratio achieved using the above algorithm depends on the wavelet function. For noise removal, we make use of 'visual analysis' technique. The noise on the channel is modeled and anlaysed using wavelet packets. The noise energy distribution is calculated in the various subspaces and at various levels. The 'subpace-energy' plot so obtained gives a clear picture of noise distribution in various subspaces. The 'noisy signal' received on this channel is analysed in the similar fashion and a similar plot is obtained. The subspaces, which correspond to high noise energy and low signal energy can be discriminated visually . To the rest of the subspaces we apply a threshold and drop all the coefficients below that threshold. The above threshold is calculated using 'hypothesis testing' technique. The noise removal upto 75 percent has been achieved.

The wavelet packet decomposition of an image has also been used for edge enhancement and edge detection by rescaling 'average signal' or 'detail signal' depending on the requirement.

Contents

1 Introduction 1

1.1 Importance of Feature Extraction 1

1.2 Overview of The Thesis 2

2 Review of Multiresolution Analysis 4

2.1 The Basic Idea 4

2.2 Introduction to Library of Orthonormal Bases 8

2.3 Wavelet Bases 9

2.4 Wavelet Packet Bases 12

2.5 Selection of a "Best Basis" from a Library of Orthonormal Bases . . . 12

2.5.1 Information Cost Function 12

2.5.2 Algorithm for Best-Basis Selection 14

2.6 Mallat's Decomposition Algorithm for 1-D Signal 15

2.6.1 Average Signal 15

2.6.2 Detail Signal 16

2.7 Mallat's Reconstruction Algorithm for 1-D Signal 16

2.8 Application of Wavelet Packets to Images 17

2.9 2-Dimensional Operators 18

2.10 Two-dimensional Wavelet Packets 21

2.11 Mallat's Decomposition Algorithm for 2-D Signal 22

2.12 Mallat's Reconstruction Algorithm for 2-D Signal 25

3 Application of Wavelet Packets in Feature Extraction 26

3.1 Development of 2-D Wavelet Package for Analysis of Images 28
3.2 Image Compression 28
3.2.1 Selection of Best Basis in Library 29
3.2.2 Compressing a Best-Basis Representation 30
3.2.3 Efficient Coding of Side Information which Describes Best-Basis 31
3.3 Reconstruction 35
3.4 Complexity of the Algorithm 35
3.5 Results of Image Compression 36
3.5.1 Determination of Compression Factor 36
3.5.2 Results 36
3.6 Edge Charecterisation 40
3.7 Noise Removal 41
3.7.1 Noise Removal using Subspace Energy Display 41
3.7.2 Use of Hypothesis Testing Technique for Calculation of Threshold 42
3.7.3 Results of Noise Removal 43
3.8 Local Image Rescaling 47
3.9 Quality Criteria for Wavelets used in Image Processing 47
3.9.1 Wavelet Regularity 47
3.9.2 Number of Vanishing Moments 48
3.9.3 Spatial Charecterization of the Scaling Function 49

4 Conclusion and Suggestions for Future Work **56**
4.1 Conclusion 56
4.2 Suggestions for Future Work 57

List of Figures

2.1 A decomposition of $\Omega_{0,0}$ into the mutually orthogonal subspaces using wavelet transform (with J=3) The symbols in bold represent the subspaces kept intact by the wavelet transform 11

2.2 A decomposition of $\Omega_{0,0}$ into the tree-structured subspaces using wavelet packet transform (with J=3). 13

2.3 Generation of detail signal at j^{th} step. 17

2.4 Reconstruction of 1-D Signal 17

2.5 Block Diagram for Convolution-Deconvolution 18

2.6 Decomposition of an Image into its First generation *descendents* . . . 20

2.7 Decomposition of an Image into Two generation *descendents* 20

2.8 Decomposition of an Image into *L levels* 21

2.9 Schematic representation of octave-subbands 22

2.10 Decomposition of a two dimensional signal down to level 4 23

2.11 Detail Signal at j^{th} resolution for 2-D image. 24

2.12 Average Signal at j^{th} step for 2-D image. 24

2.13 Reconstruction of 2-D Signal 25

3.1 Scaling functions and Wavelets used in this thesis 27

3.2 The subspaces array stacked over one another 32

3.3 Plot of Compression achieved Vs Mean Square Error for various wavelets. 37

3.4 Effect of denoising on ECG signal 44

3.5 Effect of denoising on Sine Wave signal 45

3.6 Effect of denoising on transient 46

Chapter 1

Introduction

1.1 Importance of Feature Extraction

In analyzing and interpreting signals such as musical recordings, seismic signals, or stock market fluctuations , or images such as mammograms or satellite images,extracting relevant features from them is of vital importance. Often, the important features of signal analysis, such as edges, spikes or transients, are characterized by local information either in the time (or space)domain or in the frequency (or wave number)or both: for example, to discriminate seismic signals caused by natural earthquakes,the frequency charecterstics of primary waves,which arrive in a short and specific time window,may be a key factor;to distinguish benign and malignant tissues in mammograms, the sharpness of the edges of masses may be of critical importance.

In this thesis we explore how to extract relevant features from signals and discard irrelevant information for a variety of problems in signal analysis. The term feature extraction covers following aspects of signal analysis:

Compression : how to represent and describe signals in compact manner for information transmission and storage.

Noise removal : how to remove random noise or undesired components from signals(also called discard).

Classification : how to classify signals into categories.

Regression : how to predict a response of interest from input signals.

Edge Characerization : how to detect singularities(e.g.,spikes or edges)in signals and characterize them.

In the present thesis, the work has been restricted to compression, noise removal, and edge detection.

Various methods have been used for the purposes of feature extraction. But it seems quite logical that, as far as images are concerned, the method that closely mimics the human visual system (HVS), should give comparatively better results. Use of wavelets fall in this category. In this thesis, we have gone a step further and have used wavelet packets for the same purpose.

The methods we develop here can be applied to many different types of signals and images. The signals and images treated in this thesis are all discrete: they are simply matrices consisting of finite number of real valued samples. These signals normally have very large number of samples.Therefore,extracting only important features for the problem at hand and discarding irrelevant information(this is called reduction of dimensionality) are crucial.

1.2 Overview of The Thesis

This thesis explores some of the feature extraction and signal analysis problems using the wavelet packets and best-basis paradigm. The emphasis here has been on compression and noise removal. The thesis is organised as follows.

Chapter-1 gives a brief introduction of the entire work. It introduces the term "Local Feature Extraction" and also covers the various aspects covered by this term. It also covers, in brief, the importance of feature extraction.

Chapter-2 reveiws the concept of multiresolution analysis. It illustrates the Mallat's decomposition and reconstruction algorithm. It also discusses reconstruction error involved in decomposition and reconstruction using the above algorithm.

It also discusses the formation of a library of orthonormal bases using wavelet and wavelet packet bases. It introduces terms like, library, dictionary and discusses Information cost function and Best-basis selection using information cost function. Basiscally cost function used for best basis selection is entropy.

In this chapter we also discuss two dimension wavelet packets, notion of entropy while selecting best basis and application of best basis criterion to images.

Chapter-3 discusses the use of wavelet packets and best basis in image compression, noise removal, edge detection and local image rescaling. In this chapter we also discuss the results achieved.

Chapter-4 concludes the thesis with some suggestions for the future work.

Chapter 2

Review of Multiresolution Analysis

2.1 The Basic Idea

Let $A_{2^j} f(x)$ be the operator which approximates a signal at a resolution 2^j.We suppose that our original signal $f(x)$ is measurable and has finite energy :$f(x) \in L^2(R)$. Here we characterize A_{2^j} from the intutive properties that one would expect from such an approximation operator.

1. A_{2^j} is a linear opertor If $A_{2^j} f(x)$ is the approximation of some function $f(x)$ at the resolution 2^j,then $A_{2^j} f(x)$ is not modified if we approximate it again at the resolution 2^j. This principle shows that $A_{2^j} \circ A_{2^j} = A_{2^j}$. The operator A_{2^j} , is thus a projection operator on a particular vector space $V_{2^j} \subset L^2(R)$.The vector space V_{2^j} can be interpreted as the set of all possible approximations at the resolution 2^j of functions in $L^2(R)$.

2. Among all the approximated functions at the resolution 2^j , $A_{2^j} f(x)$ is the function which is most similar to $f(x)$.

$$\forall g(x) \in V_{2^j}, \|g(x) - f(x)\| \geq \|A_{2^j} f(x) - f(x)\|. \tag{2.1}$$

Hence, the operator A_{2^j} is an orthogonal projection on the vector space V_{2^j}.

3. Containment The approximation of a signal at a resolution 2^{j+1} contains all the necessary information to compute the same signal at a smaller resolution. This is a causality property. Since A_{2^j} is a projection operator on V_{2^j} this principle is equivalent to

$$\forall j \in Z, V_{2^j} \subset V_{2^{j+1}} \tag{2.2}$$

4. An approximation operation is similar at all resolutions The spaces approximated should be thus derived from one another by scaling each approximated function by the ratio of their resolution values

$$\forall j \in Z, f(x) \in V_{2^j} \Leftrightarrow f(2x) \in V_{2^{j+1}}. \tag{2.3}$$

5. The approximation A_{2^j} of a signal $f(x)$ can be charecterized by 2^j samples per length unit When $f(x)$ is translated by a length proportional to 2^{-j}, $A_{2^j} f(x)$ is translated by the same amount and ischarecterized by the same samples which have been translated. As a consequence of (3) , it is sufficient to express this principle for the resolution $j = 0$. The mathematical translations consist of the following

- Discrete charecterization. There exists an isomorphism I from V_1 onto $l^2(Z)$.
- Translation of the approximation

$$\forall k \in Z, A_1 f_k(x) = A_1 f(x-k), \text{ where } f_k(x) = f(x-k) \tag{2.4}$$

- Translaton of samples:

$$I(A_1(f(x)) = (\alpha_i)_{i \in Z} \Leftrightarrow I(A_1(f_k(x)) = (\alpha_{i-k})_{i \in Z} \tag{2.5}$$

6. When computing an approximation of $f(x)$ at resolution 2^j some information about $f(x)$ s lost. However as the resolution increases to $+\infty$ the approximated signal should converge to the original signal. Conversely as the resolution deceases to zero, the approximated signal contains less and less information and converges to zero.

Since the approximated signal at a resolution 2^j is equal to the orthogonal projection on a space V_{2^j} , this principle can be written as

$$\lim_{j \to +\infty} V_{2^j} = \bigcup_{j=-\infty}^{+\infty} V_{2^j} \text{ is dense in } L^2(R) \tag{2.6}$$

and

$$\lim_{j \to -\infty} V_{2^j} = \bigcap_{j=-\infty}^{+\infty} V_{2^j} = \{0\} \tag{2.7}$$

We call any set of vector spaces which satisfies the above properties a *multiresolution approximation* of $L^2(R)$.

It has been seen that the approximation operator A_{2^j} is an orthogonal projection on the vector space V_{2^j}. In order to numerically charecterize this operator, we must find an orthonormal basis of V_{2^j}. Theorem (2.1) shows that such an orthonormal basis can be defined by dialating and translating a unique function $\phi(x)$.

Theorem 2.1 Let $(V_{2^j})_{j \in Z}$ be a multiresolution approximation of $L^2(R)$. There exists a unique function $\phi(x) \in L^2(R)$, called *scaling function*, such that if we set $\phi_{2^j}(x) = 2^j \phi(2^j x)$ for $j \in Z$ (the dilation of $\phi(x)$ by 2^j),then

$$\left(\sqrt{2^{-j}} \phi_{2^j}(x - 2^{-j} n)\right)_{n \in Z} \text{ is an orthonormal basis of } V_{2^j}. \tag{2.8}$$

The proof of this theorem can be found in [1].The theorem shows that an orthonormal basis of any V_{2^j} can be built by dialating a function $\phi(x)$ with a coefficient 2^j and translating a the resulting function on a grid whose interval is proportional to 2^{-j}. The functions $\phi_{2^j}(x)$ are normalised with respect to the $L^1(R)$ norm. The coefficient $\sqrt{2^{-j}}$ appears in the basis set in order to normalize the functions in order to normalize the functions in the $L^2(R)$ norm. For a given multiresolution approximation $(V_{2^j})_{j \in Z}$, there exists a unique scaling function $\phi(x)$ which satisfies the above equation.However, for different multiresolution approximation scaling functions are different.

The orthogonal projection on V_{2^j} can now be computed by decomposing the signal $f(x)$ on the orthonormal basis given by Theorem (2.1). That is,

$$\forall f(x) \in L^2(R), A_{2^j} f(x) = 2^{-j} \sum_{n=-\infty}^{+\infty} \langle f(u), \phi_{2^j}(u - 2^{-j} n) \rangle \phi_{2^j}(x - 2^{-j} n). \tag{2.9}$$

The algorithm for determining *discrete approximation* fo the signal $f(x)$ at resolution 2^j is disussed im Mallat's decomposition algorithm in the next section. Since $\phi(x)$ is a low-pass filter, this discrete signal can be interpreted as a low-pass filtering of $f(x)$ followed by a uniform sampling at the rate 2^j. The scaling function $\phi(x)$ forms a very particular low-pass flter since the family of functions $\left(\sqrt{2^{-j}}\phi_{2^j}(x-2^{-j}n)\right)_{n\in Z}$ is an orthonormal family.

Before we consider the wavelet representation, one more theorem which will be used later on is mentioned below(the proof of the same is given in [1]):

Theorem 2.2 Let $\phi(x)$ be a scaling function, and let H be a discrete filter with impulse response $h(n) = \langle \phi_{2^{-1}}(u), \phi(u-n)\rangle$ Let $H(\omega)$ be the Fourier Series defined by

$$H(\omega) = \sum_{n=-\infty}^{+\infty} h(n)e^{-\imath n\omega}. \tag{2.10}$$

$H(\omega)$ satisfies the following two properties:

$$\mid H(0) \mid = 1 \text{ and} h(n) = O(n^{-2}) \text{ at infinity} \tag{2.11}$$

$$\mid H(\omega) \mid + \mid H(\omega+\pi) \mid^2 = 1 \tag{2.12}$$

Conversely let $H(\omega)$ be a Fourier series satisfying above equation and such that

$$\mid H(\omega) \mid \neq 0 \text{ for } \omega \in [0, \pi/2] \tag{2.13}$$

The function defined by

$$\hat{\phi}(\omega) = \prod_{p=1}^{+\infty} H(2^{-p}\omega) \tag{2.14}$$

is the Fourier transform of scaling function. In he next section we will see the wavelet representation of MRA.

Our aim is to build a multiresolution representation based on the differences of information available at two succesive resolutions 2^{j+1} and 2^j. This difference of information is called the *detail signal* at the resolution 2^j. The approximation at the resolution 2^{j+1} and 2^j of the signal are respectively equal to its orthogonal projection on $V_{2^{j+1}}$ and V_{2^j}. By applying the projection theorem , we can easily

show that the detail signal at the resolution 2^j is given by the orthogonal projection of the original signal on the orthogonal complement of V_{2^j} in $V_{2^{j+1}}$. Let O_{2^j} be this complement,i.e. O_{2^j} is orthogonal to V_{2^j},

$$O_{2^j} \oplus V_{2^j} = V_{2^{j+1}}.$$

To compute the orthogonal projection of a function $f(x)$ on O_{2^j}, we need to find an orthonormal basis of O_{2^j}.Theorem (2.3) shows that such a basis can be built by scaling and translating a function $\psi(x)$.

Theorem 2.3 Let $(V_{2^j})_{j\in Z}$ be a multiresolution vector space sequence, $\phi(x)$ the scaling function, and H the corresponding conjogate filter. Let $\psi(x)$ be a function whose Fourier transform is given by

$$\hat{\psi}(\omega) = G\left(\frac{\omega}{2}\right)\hat{\psi}\left(\frac{\omega}{2}\right) \text{ with } G(\omega) = e^{-i\omega}\overline{H(\omega+\pi)} \tag{2.15}$$

Let $\psi_{2^j}(x) = 2^j\psi(2^j x)$ denote the dilation of $\psi(x)$ by 2^j then $\left(\sqrt{2^{-j}}\psi_{2^j}(x-2^{-j}n)\right)_{(n,j)\in Z^2}$ is an orthonormal basis of O_{2^j} and $\left(\sqrt{2^{-j}}\psi_{2^j}(x-2^{-j}n)\right)_{(n,j)\in Z^2}$ is an orthonormal basis of $L^2(R)$.

$\psi(x)$ is called an *orthogonal wavelet.*

2.2 Introduction to Library of Orthonormal Bases

Having reveiwed the concepts of MRA now we will look into the Library of orthonormal bases thus formed. We will also reveiw "best-basis" algorithm of Coifman and Wickerhauser which was developed mainly for signal compression and compare with the well known Karhuen-Loeve basis .

Throughout this thesis, we consider real-valued discrete signals with finite length $n(=2^{n_0})$.To focus our attention on our main theme, we assume the periodic boundry condition on the signals;a signal $x = (x_k)_{k=0}^{n-1}$ is extended periodically beyond the interval [0,n-1] with $x_k = x_k \pmod n$ for any k $\in Z$ if necessary. If one is concerned with the discontinuies created by the periodic boundry condition,their effects can be reduced by considering an evenly-folded version $x' = (x_0, \ldots, x_{n-1}, x_{n-1}, \ldots, x_0)$of period 2n.

2.3 Wavelet Bases

The wavelet transform can be considered as a smooth partition of frequency axis. The signal is first decomposed into low and high frquency bands by the convolution -subsampling operations with the pair consisting of "lowpass" filter $\{h_k\}_{k=0}^{L-1}$ and a "highpass" filter $\{g_k\}_{k=0}^{L-1}$ directly on the discrete time domain. Let $f = \{f_k\}_{k=0}^{L-1}$ be a real valued vector of even length K . Let H and G be the convolution-subsampling operators using these filters which are defined as :

$$(Hf)_k = \sum_{l=0}^{L-1} h_l f_{2k+l} (Gf)_k = \sum_{l=0}^{L-1} g_l f_{2k+l} \tag{2.16}$$

for $k = 0, 1, \ldots K-1$. Because of the periodic boundry condition on f(whose period is K), the filtered sequences Hf and Gf are also periodic with period K/2. Their adjoint operations(i.e.,upsampling-anticonvolution)H^* and G^* are defined as

$$(H^*f)_k = \sum_{0 \le k-2l < L} h_{k-2l} f_l \qquad (G^*f)_k = \sum_{0 \le k-2l < L} g_{k-2l} f_l \tag{2.17}$$

for $k = 0, 1, \ldots 2K-1$. The operators H and G are called (perfect reconstruction) quadrature mirror filter (QMFs)if they satisfy the following orthogonality (or perfect reconstruction)conditions:

$HG^* = GH^* = 0$, and $H * H + G * G = I$;

where I is is identity operator. These conditions impose some restrictions on the filter coefficients $\{h_k\}$ and $\{g_k\}$.Let m_0 and m_1 be the bounded periodic function defined by

$$m_0(\xi) = \sum_{k=0}^{L-1} h_k e^{\imath k\xi}, m_1(\xi) = \sum_{k=0}^{L-1} g_k e^{\imath k\xi} \tag{2.18}$$

Daubechies proved that H and G are QMFs if and only if the following matrix is unitary for all $\xi \in \Re$:

$$\begin{pmatrix} m_0(\xi) & m_0(\xi+\pi) \\ m_1(\xi) & m_1(\xi+\pi) \end{pmatrix}$$

Various design criteria (concerning regularity, symmetry etc.) on the lowpass filter coefficients h_k can be found. Once h_k is fixed,we can have QMFs by setting $g_k = (-1)^k h_{L-1-k}$.

This decomposition(or expansion or analysis) process is iterated on the low bands and each time the high frequency coefficients are retained intact. At the last iteration , both, low and high frequency coefficients are kept. In other words ,let x= $\{x_k\}_{k=0}^{n-1} \in \Re^n$ be a vector to be expanded. Then , the convolution -subsampling operation transform the vector x into two subspaces Hx and Gx of lengths n/2. Next, the same operations are applied to the vector of lower frequency band to $Hx$to obtain H^2x and GHx of lengths n/4.If the process is iterated $J(\leq n_0)$ times , we have the discrete wavelet coefficients$(Gx, GHx, GH^2x, \ldots GH^Jx, H^{J+1}x)$of length n.As a result , the wavelet transform analyzes the data by partitioning its frequency content dyadically finer and finer toward the low frequency region(i.e ,coarser and coarser in the original time domain).

If we were to partition the frequency axis sharply using the charecter using the characterstic functions(or box-carfunctions),then we would have ended up the so called Shanon(or Littlewood-Paley) wavelets, i.e., the difference of two sinc functions. Clearly, however,we cannot have a finite length filter in the time time domain in this case. The other extreme is Haar basis which partitions the frequency axis quite badly but gives theshortest filter length.($L = 2$ with $h_0 = h_1 = 1/\sqrt{2}$)in the time domain and which are suitable for describing discontinuous functions.

The reconstruction (or synthesis)process is very simple : starting from the lowest two frequency bands $\mathbf{H^{J+1}x}$ and $\mathbf{GH^Jx}$, the adjoint operators are applied and added to obtain $\mathbf{H^Jx = H^*H^{J+1}x + G^*GH^Jx}$. This process is iterated to reconstruct original vector x. The computational complexity of the decomposition and reconstruction process is in both cases as easily seen.

Because of the perfect reconstruction condition on H and G, each decomposition step is also considered as a decomposition of vector space into mutually orthogonal subspaces. For the purpose of explanation we will now follow a different notation

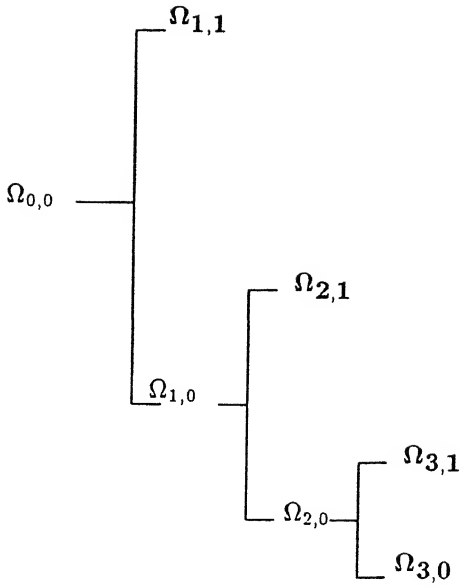


Figure 2.1: A decomposition of $\Omega_{0,0}$ into the mutually orthogonal subspaces using wavelet transform (with J=3).The symbols in bold represent the subspaces kept intact by the wavelet transform

to identify a subspace. Let $\Omega_{0,0}$ denote the standard vector space R^n. Let $\Omega_{1,0}$ and $\Omega_{1,1}$ be mutually orthogonal subspaces generated by application of of the projection operators H and G respectively to parent space $\Omega_{0,0}$.Then, in general, jth step of decomposition process can be written as:

$$\Omega_{j,0} = \Omega_{j+1,1} \oplus \Omega_{j+1,1} \text{ for } j = 0, 1, \ldots J.$$

It is clear that $\dim\Omega_{j,} = 2^{n_0-j}$. The wavelet transform is simply a way to represent $\Omega_{0,0}$ by a direct sum of mutually orthogonal subspaces.

$$\Omega_{0,0} = \left(\bigoplus_{j=1}^{J} \Omega_{j,1}\right) \oplus \Omega_{J,0}, \tag{2.19}$$

and the decomposition process is illustrated by Fig (2.1)

2.4 Wavelet Packet Bases

For oscillating signals such as acoustics signals, however , the analysis by wavelet transform is not always efficient because it only partitions the frequency axis finely towards the low frequency. The *wavelet packet transform* is a generalised version of the wavelet transform: it decomposes even the high frequency bands which are kept intact in wavelet transform. They are much more oscillatory in nature as compared to the wavelet basis vectors. The first level of decomposition generates Hx and Gx just like in wavelet transform. The second level generates four subsequences,H^2x, GHx, HGx, G^2x. If we repeat this process for J times, we end up having $J n$expansion coefficients. It is easily seen that the computational cost of this whole process is about $O(Jn) \leq O(n \log_2 n)$ This iterative process naturally generates subspaces with different frequency localizaion characteristics. The root node of this tree is again $\Omega_{0,0}$. The node $\Omega_{j,k}$ splits into the two orthogonal subspaces $\Omega_{j+1,2k}$ and $\Omega_{j+1,2k+1}$ by the operators H and G , respectively:

$$\Omega_{j,k} = \Omega_{j+1,2k} \oplus \Omega_{j+1,2k+1} for j = 0, 1, \ldots J, k = 0, \ldots 2^j - 1.$$

The Fig (2.2) shows the binary tree of subspaces of $\Omega_{0,0}$:

Clearly, we have redundant set of subspaces in the binary tree. In fact ,it is easily proved that there are more than $2^{2^{J-1}}$ possible orthonormal bases in the binary tree. This binary tree (or as we see later the quad tree in case of images) is one of the most important tools.

2.5 Selection of a "Best Basis" from a Library of Orthonormal Bases

2.5.1 Information Cost Function

An efficient coordinate system for representing signal should give large magnitudes along a few axes and negligible magnitudes along most axes when the signal is expanded into the associated basis. We then need a measure to evaluate and compare

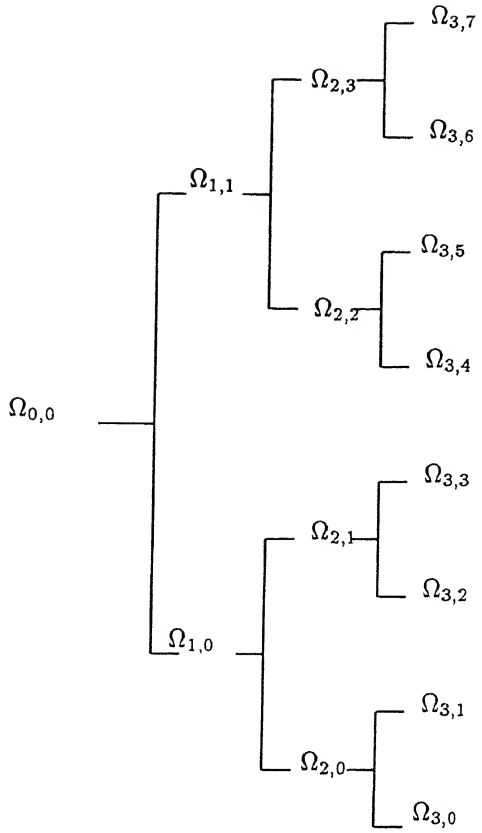


Figure 2.2: A decomposition of $\Omega_{0,0}$ into the tree-structured subspaces using wavelet packet transform (with J=3).

the efficiency of many bases. Let $\mathcal{I}$ denote this measure which is often called "information cost" function. There are several choices for $\mathcal{I}$. All of them essentially measure the "energy concentraton " of coordinate vector. A natural choice for this measure is the *Shanon entropy* of the coordinate vector. Let us define the entropy of a nonnegative sequence $\mathbf{p} = \{p_i\} with\ \sum_i p_i = 1$ by

$$H(\mathbf{p}) \underline{\underline{\Delta}} - \sum_i p_i \log_2 p_i \ldots \tag{2.20}$$

with the convention $0 \cdot \log 0 = 0$ (From now on, we use "log" for the logarithm of base 2.) For a signal $\mathbf{x}$,we set $p_i = (\mid x_i \mid / \|\mathbf{x}\|_r)^r$ where $\|\cdot\|_r$ is the ℓ^r norm and $1 \leq r < \infty$ and define

$$H_r(\mathbf{x}) \underline{\underline{\Delta}} - \sum_i \frac{\mid x_i \mid^r}{\|\mathbf{x}\|^r} \log \frac{\mid x_i \mid^r}{\|\mathbf{x}\|^r} \ldots \tag{2.21}$$

Often $r = 1$ or $r = 2$ is used. In this thesis we will use $r = 2$.

2.5.2 Algorithm for Best-Basis Selection

The "best-basis" algorithm of Coifman and Wickerhauser(this will be discussed later in details) was developed mainly for signal compression. This method first expands a given *signal* into a specified library of orthonormal bases. Then a complete basis called a *best basis*(**BB**) which minimizes a certain information cost function such as entropy is searched int this binary tree using devide-and-conquer algorithm.

$$B_{j,k} = \left(\omega_{j,k,0}, \ldots \omega_{j,k,2^{n_0-j}-1}\right)^T \ldots \tag{2.22}$$

Algorithm 2.1 (The Best-Basis Algorithm) *Given a vector x*

Step 1 *Choose a library of orthonormal bases $\mathcal{D}$ (i.e. specify QMFs for a wavelet packet dictionary) and specify the maximum depth of decomposition J and an information cost $\mathcal{I}$*

Step 2 *Expand* $\mathbf{x}$ *into the dictionary* $\mathcal{D}$ *and obtain coefficients* $\{B_{j,k}\mathbf{x}\}_{0\le j<J,\, 0\le k<2^j-1}$.

Step 3 *Set* $A_{J,k} = B_{J,k}$ *for* $k = 0, \ldots 2^J - 1$.

Step 4 *Determine the best subspace* $A_{j,k}$ *for* $j = J-1, \ldots 0$, $k = 0, \ldots 2^j - 1$ *by*

$$A_{j,k} = \left\{ \begin{array}{ll} B_{j,k} & \text{if } \mathcal{I}(B_{j,k}\mathbf{x}) \le \mathcal{I}(A_{j+1,2k}\mathbf{x} \cup A_{j+1,2k+1}\mathbf{x}i) \\ A_{j+1,2k} \oplus A_{j+1,2k+1} & \text{otherwise.} \end{array} \right\} \tag{2.23}$$

To make this algorithm fast, the cost function $\mathcal{I}$ needs to be *additive*.

2.6 Mallat's Decomposition Algorithm for 1-D Signal

We have seen that as $\phi_{j,k}, k \in Z$ is the orthonormal basis for V_{2^j}, any function belonging to $L^2(R)$ can now be decomposed along all translates of $\phi(2^j x)$. The result is denoted by $A^d_{2^j}$. Similarly, the function can be decomposed along all translates of $\psi(2^j x)$ and the result is denoted by $D^d_{2^j}$. If the actual signal is having N number of samples these results are having $N/2^j$ number of samples. Thus the resulting signals belong to $L^2(N/2^j)$. This is illustrated for one dimensional case and then extended to two dimensional case.

The decomposed signal along all the translates of $\phi(2^j x)$ is called the averaging signal and the decomposed signal along all translates of $\psi(2^j x)$ is called the detail signal. This is due to the nature of finite sequences $h(n)$ and $g(n)$ used to generate the $\phi(x)$ and $\psi(x)$ respectively.

2.6.1 Average Signal

It is the inner product between the given function and translates of $\phi(2^j x)$. It is denoted by

$$\langle f(x), \sqrt{2^j}\phi(2^j x - k)\rangle, \ \forall k \in Z$$

The discrete version is given by

$$\langle f(x_i), \sqrt{2^j}\phi(2^j x_i - k)\rangle, \ \forall k \in Z \text{ and } x_i \text{ is position of } i^{th} \text{ sample.}$$

Equivalently, the inner product is written as

$$\begin{aligned} &= \sum_{x_i} f(x_i)\sqrt{2^j}\phi(2^j x_i - k) \\ &= \left(f(x_i) * \sqrt{2^j}\phi(-2^j x_i)\right)(2^{-j}k), \ k \in Z \end{aligned} \tag{2.24}$$

By changing the variables and denoting $h(j) = \langle \phi_{2^{-1}}(x_i), \phi(x_i - j)\rangle$ and $h^*(j) = h(-j)$ the Eqn (2.24) is written as

$$\begin{aligned} &= \sum h^*(2x_i - k)\left(f(x_i) * \sqrt{2^j}\phi(-2^{j+1}x_i)\right)(2^{-j-1}k), k \in Z \\ &= A^d_{2^j}(f). \end{aligned} \tag{2.25}$$

where x_i is the location of i^{th} sample.

This implies that it is the average signal at the j^{th} resolution level obtained through convolution and decimation operation between the previous average signal $(A_{2^j}^d + 1)f$ and the h^* sequence .

2.6.2 Detail Signal

The detail signal at the j^{th} resolution step is the inner product between the given signal and the translates fo $\psi(2^j x)$. It is denoted by

$$\langle f(x), \sqrt{2^j}\psi(2^j x - k)\rangle, \forall k \in Z$$

The discrete version is given by

$\langle f(x_i), \sqrt{2^j}\psi(2^j x_i - k)\rangle, \forall k \in Z$ and x_i is the location of the i^{th} sample of signal.

Equivalently,

$$\begin{aligned} &= \sum_{x_i} f(x_i)\sqrt{2^j}\psi(2^j x_i - k) \\ &= \left(f(x_i) * \sqrt{2^j}\psi(-2^j x_i)\right)(2^{-j}k), k \in Z. \end{aligned} \tag{2.26}$$

By changing the variables and writing $g(j) = \langle \psi_{2^{-1}}(x_i), \phi(x_i - j)\rangle$ and $g^*(j) = g(-j)$, Eqn. (2.26) can be written as,

$$\begin{aligned} &= \sum g * (2x_i - k)\left(f(x_i) * \sqrt{2^j}\psi(-2^{j+1}x_i)\right)(2^{-j-1}k), k \in N. \\ &= D_{2^j}^d(f) \end{aligned} \tag{2.27}$$

This implies that the detail signal at j^{th} step is obtained through convolution and decimation operation between the previous signal $(A_{2^j}^d + 1)f$ and the g^* sequence as given in Fig (2.3).

2.7 Mallat's Reconstruction Algorithm for 1-D Signal

As we have seen earlier that

$$V_{2^{j+1}} = V_{2^j} \oplus O_{2^j}.$$

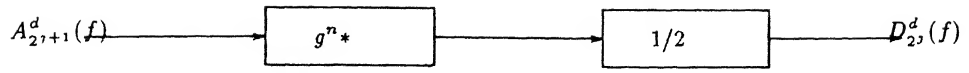


Figure 2.3: Generation of detail signal at j^{th} step.

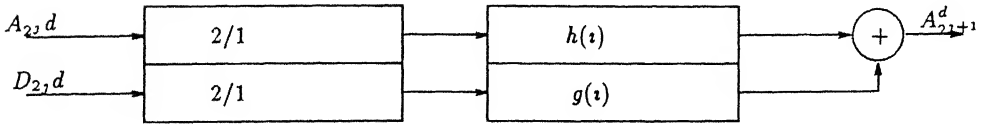


Figure 2.4: Reconstruction of 1-D Signal

So, $V_{2^{j+1}}$ can be constructed by the the orthogonal sum of approximation signal and detail sgnal in lower resolution stage. The reconstructed signal at $(j+1)$ resolution step is the inner product between the original signal and the translates of $\phi(2^{j+1}x)$. Mathematically reconstucted signal is given by,

$$\langle f(u), \sqrt{2^{j+1}}\phi(2^{j+1}u-k)\rangle \tag{2.28}$$

By changing variables and defining the h and g sequences Eqn (2.28) can be written as

$$\begin{aligned} &= \textstyle\sum_k h(n-2k)\langle f(u), \sqrt{2^j}\phi(2^j u-k)\rangle + \sum_k g(n-2k)\langle f(u), \sqrt{2^j}\psi(2^j u-k)\rangle \\ &= \textstyle\sum_k h(n-2k)A^d_{2^j} + \sum_k g(n-2k)D^d_{2^j} \end{aligned} \tag{2.29}$$

This implies that $A^d_{2^{j+1}}(f)$ can be reconstructed by putting zeros between each samples of $A^d_{2^j}(f)$ and $D^d_{2^j}(f)$ and convolving the resulting signals with filters h and g respectively. The original signal can be reconstructed by repeating this procedure down to $j=0$. This can be shown as in Fig (2.4)

2.8 Application of Wavelet Packets to Images

We have seen that a bank of perfect reconstruction QMFs can be used for the purposes of analysis of a one-dimension signal to produce a dictionary of best basis,

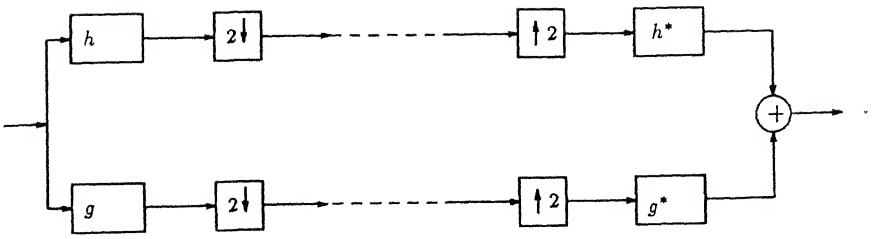


Figure 2.5: Block Diagram for Convolution-Deconvolution

as also to reconstruct the original signal from the best basis. The same concept can be extended to 2-D Images.

2.9 2-Dimensional Operators

Fig (2.5) is the traditional diagram describing the action of a pair of quadrature mirror filter.On the left is convolution and downsampling(by 2);on the right is up-sampling (by 2) and adjoint convolution, followed by summing of components. The broken lines in the middle represent either transmission or storage.
Now we can define 4 2-dimensional convolution-decimation operators in terms of H and G namely the tensor products of the pair of quadrature mirror filters:

$$\begin{aligned}
F_0 \underline{\underline{\triangle}} H \otimes H \quad & F_0 v(x,y) = \textstyle\sum_{i,j} v(i,j) h_{2x+i} h_{2y+j} \\
F_1 \underline{\underline{\triangle}} H \otimes G \quad & F_1 v(x,y) = \textstyle\sum_{i,j} v(i,j) h_{2x+i} g_{2y+j} \\
F_2 \underline{\underline{\triangle}} G \otimes H \quad & F_2 v(x,y) = \textstyle\sum_{i,j} v(i,j) g_{2x+i} h_{2y+j} \\
F_3 \underline{\underline{\triangle}} G \otimes G \quad & F_3 v(x,y) = \textstyle\sum_{i,j} v(i,j) g_{2x+i} g_{2y+j}
\end{aligned} \tag{2.30}$$

These convolution-decimation operators have the following adjoints:

$$\begin{aligned}
F_0^* v(x,y) &= \textstyle\sum_{i,j} v(i,j) h_{2i+x} h_{2j+y} \\
F_1^* v(x,y) &= \textstyle\sum_{i,j} v(i,j) h_{2i+x} g_{2j+y} \\
F_2^* v(x,y) &= \textstyle\sum_{i,j} v(i,j) g_{2i+x} h_{2j+y} \\
F_3^* v(x,y) &= \textstyle\sum_{i,j} v(i,j) g_{2i+x} g_{2j+y}
\end{aligned} \tag{2.31}$$

The orthogonality relations for the above are

$$F_n F_m^* = \delta_{nm} I$$
$$I = F_0^* F_0 \oplus F_1^* F_1 \oplus F_2^* F_2 \oplus F_3^* F_3 \tag{2.32}$$

By a picture we will mean any function $\mathbf{S} = S(x,y) \in l^2(Z^2)$. In practice , of course,pictures will be nonzero only at finitely many points. The space $l^2(Z^2)$ of pictures may be decomposed into a partially ordered set $\mathbf{W}$ of subspaces $W(n,m)$, where $m \geq 0$, and $0 \leq n < 4^m$. These are images of orthogonal projections composed of products of convolution-decimations. By putting $W(0,0) = l^2$, and defining recursively

$$W(4n+i, m+1) = F_{2i+j}^* F_{2i+j} W(n,m) \quad for\ i = 0,1,2,3 \tag{2.33}$$

These subspaces may be partially ordered by a relation which we define recursively as well. We say W is a precursor of W' (write $W \prec W'$) if they are equal or if $W' = F^* F W$ for a convolution-decimation F in the set $\{F_0, F_1, F_2, F_3\}$. We also say that $W \prec W'$ if there is a finitesequence $V_1, \ldots, V_n$ of subspaces in $\mathbf{W}$ such that $W \prec V_1 \prec \ldots \prec V_n \prec W'$. This is well defined,since each application of F^*F increases the index m.

Subspaces of a single precursor $W \in \mathbf{W}$ will be called its *descendents* ,while the first generation of descendents will naturally be called *children* . By orthogonality condition,

$$W = F_0^* F_0 W \oplus F_1^* F_1 W \oplus F_2^* F_2 W \oplus F_3^* F_3 W \tag{2.34}$$

The right hand side contains all children of W.

The coordinates with respect to the standard basis of $W(n,m)$ are in $F_{(1)} \ldots F_{(m)} W(0,0)$, where the particular filters $F_{(1)} \ldots F_{(m)}$ are determined uniquely by n. Therefore we can express in standard coordinates the orthogonal projections of $W(0,0)$onto the complete tree of subspaces W by recursively convolving and decimating with filter. We may think of the quadrature mirror filters H and G as low-pass and high-pass filters, respectively. They can be described as nominally deviding the support set

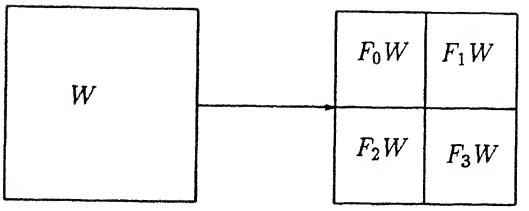


Figure 2.6: Decomposition of an Image into its First generation *descendents*

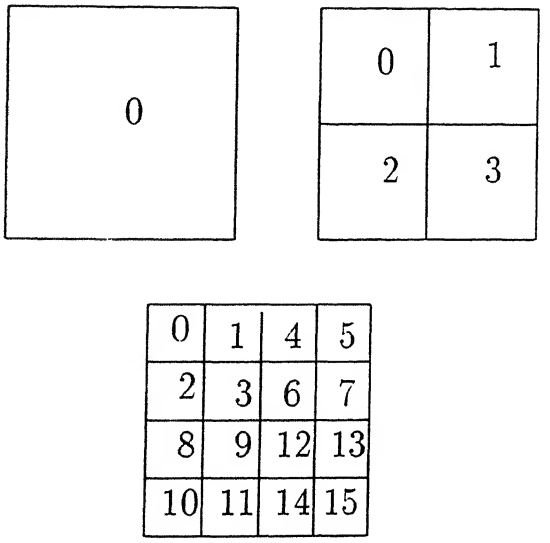


Figure 2.7: Decomposition of an Image into Two generation *descendents*

of the Fourier transform $\hat{S}$ of the picture into dyadic squares. If the filters were perfectly sharp,then this would be literally true, and the children of W would correspond to 4 dyadic subsquare onne scale smaller as in Fig (2.6)

Fig (2.7) shows two generations of descendents ,the complete decomposition of $\Re^4 \times \Re^4$. The subbands are labeled by the "n" index in $W(n,m)$. Within the dyadic squares are the n indices of the corresponding subspaces at that level. If we had started with a picture of $N \times N$ pixels, then we could repeat this decomposition process $\log_2(N)$ times.

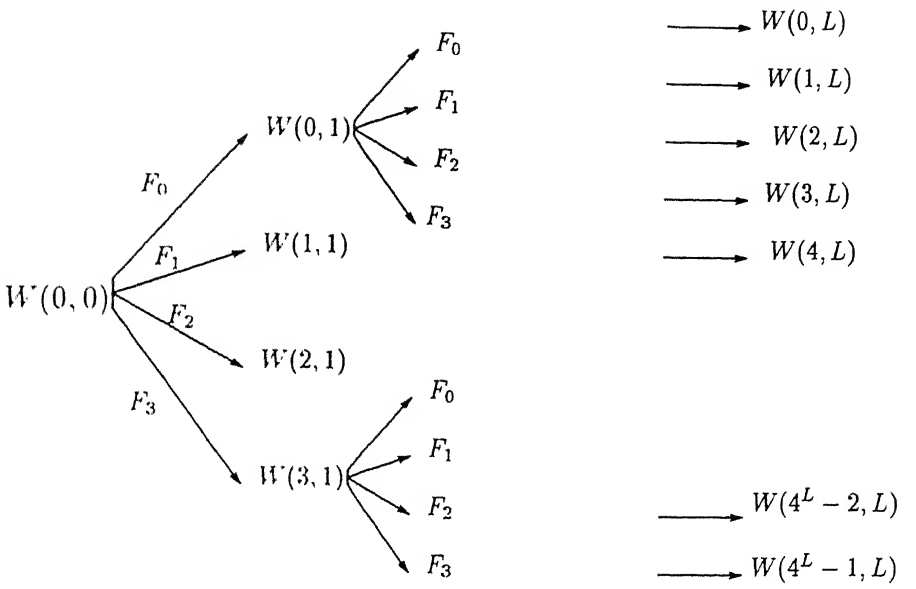


Figure 2.8: Decomposition of an Image into L *levels*

2.10 Two-dimensional Wavelet Packets

We have seen in the earlier chapter , a method for generating a library of orthonormal bases by recursive QMF convolution-decimation. From an algorithmic point of view , this is equivalent to recursive subband coding while retaining all intermediate subband decompositions. The coefficients produced at each stage are correlations of the signal with compactly-supported oscillatory functions called "wavelet packets". If perfect reconstruction QMFs are used , then the wavelet packets satisfy some remarkable orthogonal properties.

From the tree **W** of subspaces we may choose a *basis subset*,defined as a collection of mutually orthogonal subspaces $W \in \mathbf{W}$ or lists of pairs (n, m), which together span the root. Basis subsets are in one to one correspondence with dyadic decomposition of the unit square. Classical subband coding takes coefficients from a fixed set of subbands,usually from a single levelof trees as in figueres above. Wavelet transform coding also extracts coefficients from a fixed collection of blocks, the octave subbands, which are schematically represented by Fig (2.9)

Fig (2.10) shows a typical dyadic decomposition down to level 4;its basis subset

CENTRAL LIBRARY
I. I. T., KANPUR
Acc No. A 121214

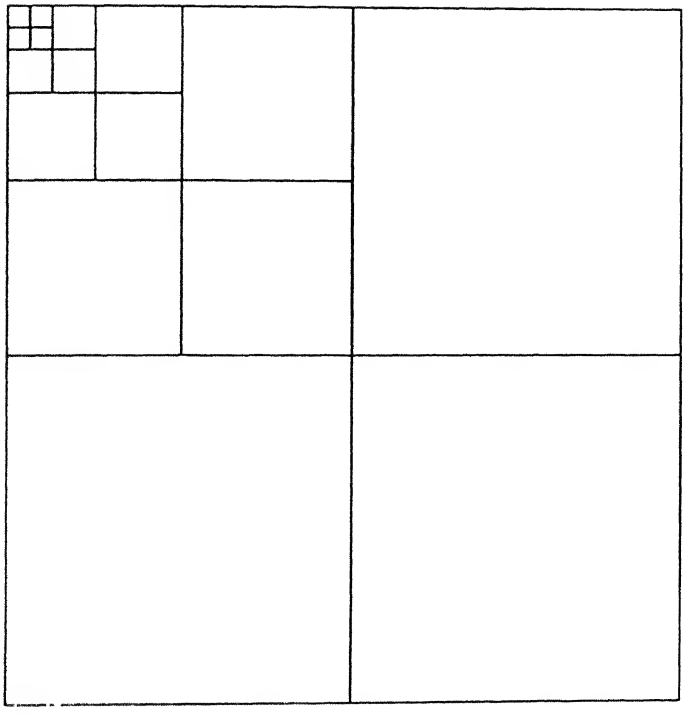


Figure 2.9: Schematic representation of octave-subbands

consists of the 16 pairs (0,1),(3,1),(4,2), (5,2),(6,2),(7,2),(8,2),(9,2),(10.2) ,(45,3),(46,3),(47,3),(176,4),(177,4),(178,4),and (179,4)

2.11 Mallat's Decomposition Algorithm for 2-D Signal

Average Signal

In this a separable scaling function is considered. The inner product between digitised input signal $f(x_i, y_i)$ and $\sqrt{2^j}\phi(2^j x_i - k)\phi(2^j y_l)$ is given by

$$\begin{aligned} &= \sum_{x_i} \sum_{y_i} f(x_i, y_i)\sqrt{2^j}\phi(2^j x_i - k)\phi(2^j y_i - l) \\ &= \left((f(x_i, y_i) * \sqrt{2^j}\psi(-2^j x_i))(2^{-j}k) * \phi(-2^j y_i)\right)(2^{-j}l), k, l \in Z \end{aligned} \tag{2.35}$$

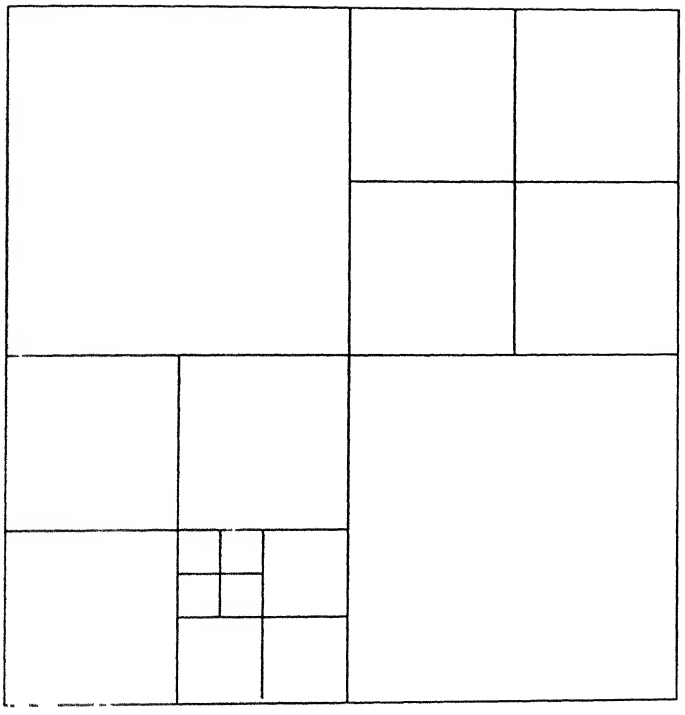


Figure 2.10: Decomposition of a two dimensional signal down to level 4

By changing variables and defining h^* sequence as in 1-D case the Eqn (2.35) can be written as,

$$= \sum_{k,l} h^*(2x_i \quad k)h^*(2y_i - l)\left(f(x_i, y_i) * \sqrt{2^j}\phi(2^{j+1}x_i) * \phi(2^{j+1}y_i)\right)(2^{-j-1}l), k, l \in N$$

$$= A^d_{2^j}(f) \tag{2.36}$$

This implies that the average signal at the j^{th} resolution step is obtained through double convolution and decimation operation between the previous average signal and $h^*(x_i)$ and $h^*(y_i)$ sequences respectively as shown in Fig (2.12).

Detail Signal

Here we consider separable scaling and mother wavelet functions. As shown in [1],[2] we note that following functions are ortogonal to each other:

- $\phi(x)\phi(y) \perp \phi(x)\psi(y)$
- $\phi(x)\phi(y) \perp \phi(y)\psi(x)$

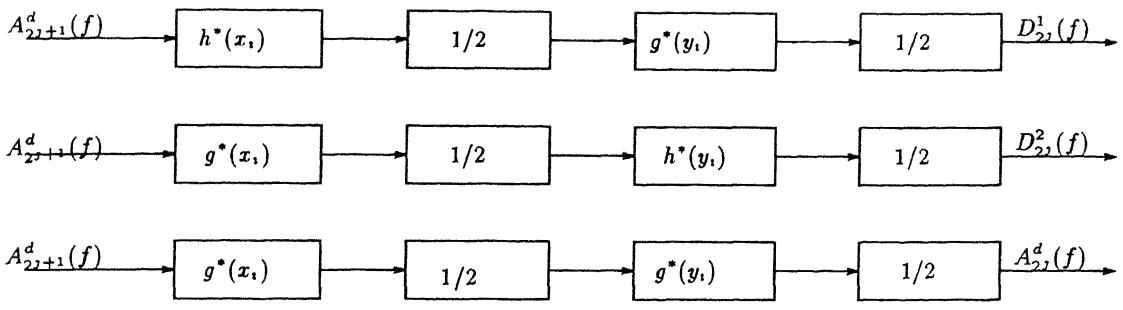


Figure 2.11: Detail Signal at j^{th} resolution for 2-D image.

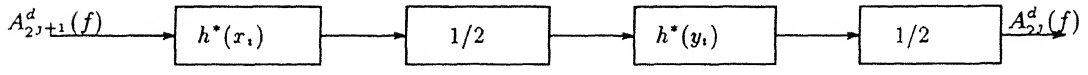


Figure 2.12: Average Signal at j^{th} step for 2-D image.

- $\phi(x)\phi(y) \perp \psi(x)\psi(y)$
- $\psi(x)\psi(y) \perp \phi(y)\psi(x)$
- $\psi(x)\psi(y) \perp \phi(x)\psi(y)$
- $\psi(x)\psi(y) \perp \phi(x)\phi(y)$

As in 1-D case we can write

$$V_{2^{j+1}} = V_{2^j} \oplus \vee_{2^1} j \oplus O_{2^2} j \oplus O_{2^3} j.$$

Where,

$O_{2^1} j$ is Vector space generated by translated functions of $\phi(2^j x)\psi(2^j y)$.

$O_{2^2} j$ is Vector space generated by translated functions of $\psi(2^j x)\phi(2^j y)$.

$O_{2^3} j$ is Vector space generated by translated functions of $\psi(2^j x)\psi(2^j y)$.

If $D^i_{2^j}$ represents the detail signal at resolution 2^j along the i^{th} direction, we have the result as shown in Fig (2.11). The approximation or average signal at the j^{th} resolution is obtained through the convolution decimation structure given in Fig (2.12).

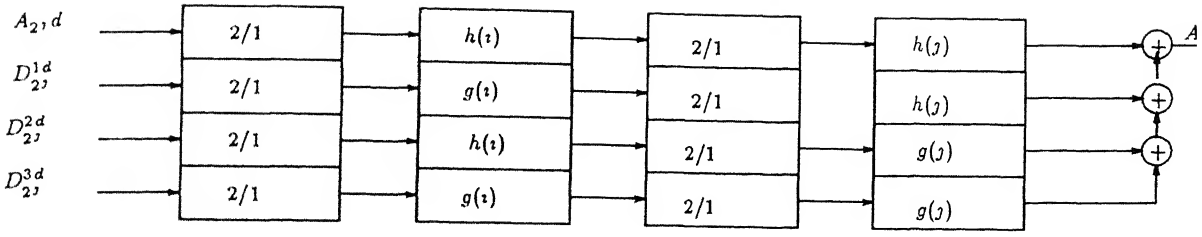


Figure 2.13: Reconstruction of 2-D Signal

2.12 Mallat's Reconstruction Algorithm for 2-D Signal

We have

$$V_{2^{J+1}} = V_{2^J} \oplus V_{2^1 J} \oplus O_{2^2 j} \oplus O_{2^3 j}.$$

Where,

$O_{2^1 j}$ is Vector space generated by translated functions of $\phi(2^J x)\psi(2^J y)$

$O_{2^2 j}$ is Vector space generated by translated functions of $\psi(2^J x)\phi(2^J y)$.

$O_{2^3 j}$ is Vector space generated by translated functions of $\psi(2^J x)\psi(2^J y)$.

If $D_{2^J}^i$ represents the detail signal at resolution 2^J along the i^{th} then by Fig (2.13) we can get the original signal.

In next chapter, we will see as to how the wavelet packets can be applied for the purpose of feature extraction in case of 2-D signals.

Chapter 3

Application of Wavelet Packets in Feature Extraction

In the previous chapter we have seen as to how we can apply wavelet packets to two dimensional signals. In this chapter we will use these two-dimensional wavelet packets for the various aspects of feature extraction as discussed in chapter1. We will consider these aspects in the subsequent sections. We have used two dimensional separable wavelets throughout the thesis work. The wavelets that have been used are:-

- Haar
- Coiflet
- Daubechies-4
- Daubechies-8
- Daubechies-20

The scaling functions ϕ and the wavelet function ψ in respect of these wavelets are given as Fig (3.1).

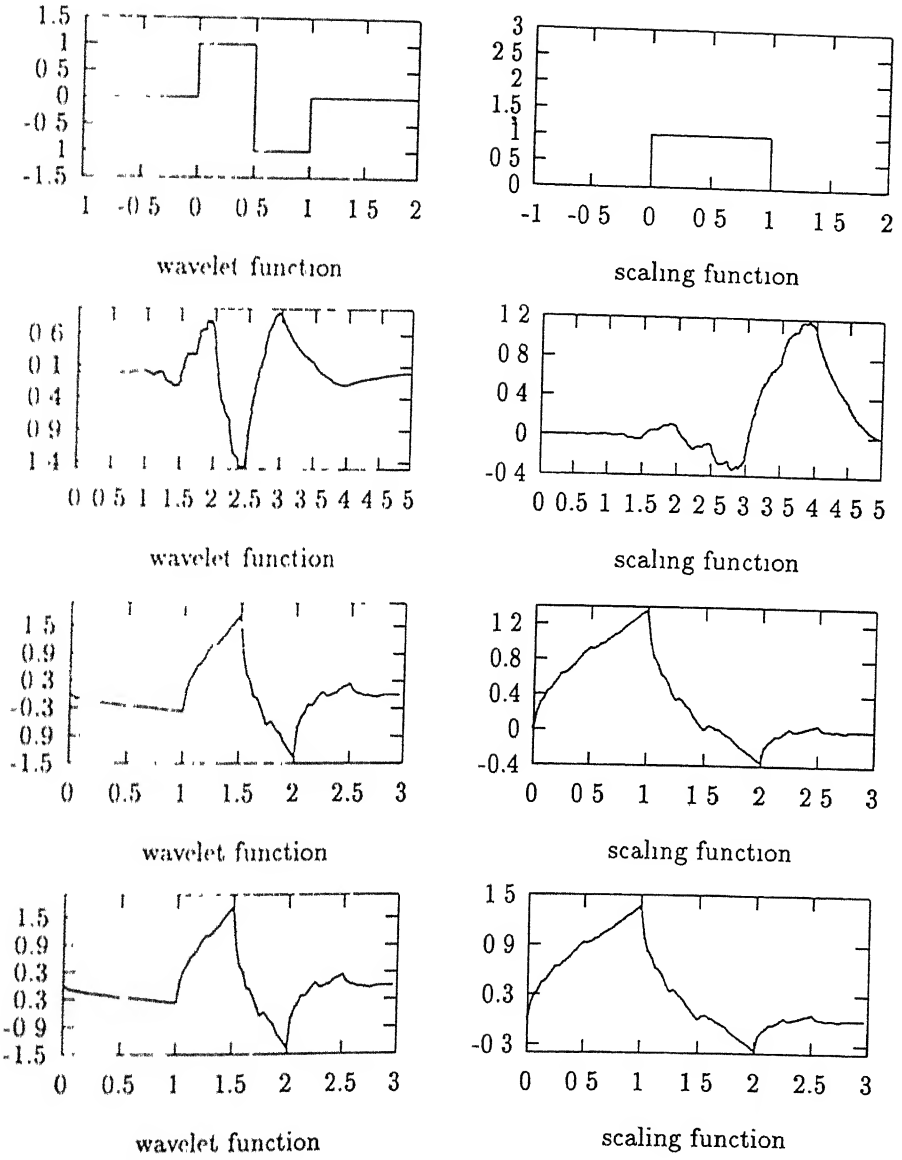


Figure 3.1: Scaling functions and Wavelets used in this thesis

3.1 Development of 2-D Wavelet Package for Analysis of Images

A package has been developed for the purposes of analysis and processing of images using wavelet packets. The features of this package are described as follows.

(a) It analyses both 1-D and 2-D signal using a variety of wavelets to generate a library of orthonormal bases. In the case of 2-D signals separable filters are used. The package also reconstructs the signal from the above library.

(b) It selects the 'best-basis' from the above library. The algorithm for selection of 'best-basis' is given in subsection (3 2.1). Once the best-basis has been selected, the package is then used to compress the same and code the additional information required to be transmitted along with the best-basis.

(c) The package also computes the energy distribution of a signal and then displays this information as shown in Plates (6) and (7).

(d) From the comparative displays of noise energy and received signal energy we can decide on the subspaces to be dropped or thrsholded as explained in subsection (3.7.1). Based on the algorithm discused in subsections (3.7.1)) and (3.7.2). The package then reconstructs the signal from the 'noisy signal'.

(e) Provisions to do edge enhancement, edge detection and rescaling of image is also included with the package developed.

3.2 Image Compression

Traditional image compression techniques have been designed to exploit the stastical redundancy present within real world images. DCT, DPCM and the entropy coding of subband images are all examples of this stastical approach. Removing redundancy can only give a limited amount of compression. To achieve high ratios some non-redundant information must be removed. The statastical coders produce

an annoying visual degradation when operating in this mode because they produce errors in visually importants parts of the image structure such as edges.

By using 'wavelets basis' for image compression some success was achieved to overcome the above problems. In this thesis we have discussed yet another approach, viz, use of wavelet packet bases for the purpose of image compression. As we will see subsequently, wavelet packets provide more flexibility and better compression. As we have discussed in the previous chapter, we use 'best-basis' algorithm for this purpose. The steps involved in this process are:-

(a) Analyse the 2-D signal using library of two dimensional wavelet packets, and form a quad-tree of depth L.

(b) From the above library, select the best-basis using some information cost function.

(c) Compress the best basis by selecting a threshold.

These processes are discussed in the subsequent sections.

3.2.1 Selection of Best Basis in Library

We have seen the algorithm for selection of best-basis in the last chapter. Now we will modify the same for a 2-D signal. Set a predetermined deepest level L. Label each subspace at level L as "kept", i.e. ,the subspaces indexed by (n, L) for $0 \leq n < 4^L$. Next, we set the level index to $L-1$. The information cost of the subspace $W(n,m)$ s then compares with the sum of the informaton costs of its children $W(4n, m+1), W(4n+1, m+1), W(4n+2, m+1), W(4n+3, m+3)$. If the parent is less than or equal to the sum of the children, then the parent is marked as "kept". This means that by choosing the parent rather than the children, we will have fewer coefficients above threshold in the representation of S. On the other hand, if the sum of children is less than the parent, we leave the parent unmarked but attribute to her the sum of children's information costs. By passing this along,

prior generations will always have their information costs compared to the least costly descendents.

After all the subspaces at level $m = L - 1$ have been compared to their children ,the level index is decremented and we continue the comparison. We can proceed this way until we have have compared the root $W(0,0)$ to its 4 children. We claim that the topmost "kept" nodes W with no kept precursors is a basis subset which minimizes the information cost. This can easily be proved by induction.

3.2.2 Compressing a Best-Basis Representation

In practice, it is necessary to compress the signal due to constraints like bandwidth. Hence, the best-basis information, determined by the algorithm discussed above has to be further compressed. Methods to achieve this compression are as follows.

Existing Method

This method has been suggested by Wickerhauser [4]. Suppose that B is a best-basis subset of W. We may then extract just non-negligible coefficients and transmit them, together with their locations in the tree. This number of coefficients is no greater than the number of pixels, since it is chosen after comparison with the original basis, among others. In that case we may use the entropy cost function to obtain the most concentrated representation, and then take only as many of the largest coeficients as we can afford. This may be accomplished by first sorting into decreasing order by absolute value,then reading off the desired number of coefficients. Alternatively,since we know in advance how many coefficients we can use,it may be more efficient to bubble up the top few coefficients and discard the rest of the array. The second method is better if the number of retained coefficients is less than $\log N$, where N is the number of pixels.

Method used in the Thesis for Compressing the Best-Basis

In the present thesis we have used energy criterion for compressing the best-basis. Once the best-basis has been selected, we calculate the normalized energy of each subspace. We then find the maximum normalized energy and select a threshold ϵ as some fraction f of this maximum energy. Then, we discard all the subspaces having energy less than this ϵ. The fraction, f has been selected based on the desired SNR and peak SNR of the image.

3.2.3 Efficient Coding of Side Information which Describes Best-Basis

The transformation from a 2-Dimensional signal to its best-basis representation is non-linear since the choice of basis depends upon the signal itself. Together with the coefficients we must include the extra information describing which basis was selected.

There are atleast 2 ways to include the information about the *basis* that were selected. Best-basis coefficients may be individually tagged with their coordinates in the best-basis tree. Suppose this tree begins with an $N \times N$ signal and decomposes it down to level L, where $L \leq \log N$. Then there are LN^2 wavelet packets coefficients, and it takes $\log 2^N$ bits to encode.

Alternatively, we may agree upon an ordering of coefficients,write out the coefficients in this order after quantization , and then entropy code the entire list. If we were to use a single basis like wavelets then we need never explicitly tag any coefficients. We obtain compression because the quantized stream of coefficients has a lower entropy than that of the original signal.

We shall now discuss second method, whereby we will include some side information which describes the chosen basis, and we shall then write all the quantized coefficients from that basis out into a stream for entropy coding. This method is substantially more efficient , and is esential for a competitive picture compression algorithm.

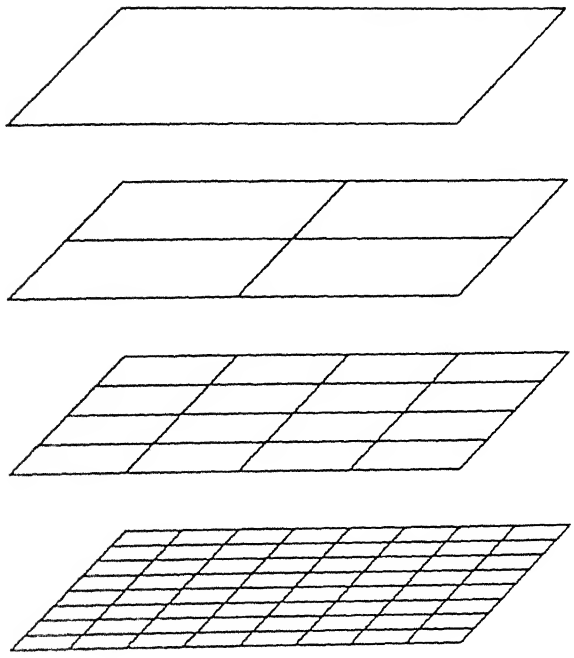


Figure 3.2: The subspaces array stacked over one another

Describing the basis quantizing all the coefficients Imagine $L+1$ arrays of $N \times N$ numbers. The first array represents the original signal, which we may call Z. the second is a concatenation of 4 subspaces obtained via sepearable filter convolution-decimation,i.e. , the spaces $F_0(X)F_0(Y)Z, F_1(X)F_0(Y)Z, F_0(X)F_1(Y)Z,$ and $F_1(X)F_1(Y)Z$. Array m represents the concatenation of the 4^m subspaces that make up level m of the wavelet packet decomposition. Ofcourse,we must satisfy the condition $0 \leq m \leq L \leq \log N$. These arrays can be visualised as stacked one atop other (as shown Fig (3.2)).

Suppose that from this collection of arrays we have chosen a best-basis. This will be a subset of the coefficients having the property that if one element of a subspace is in the basis , then whole subspace is in the basis. Such a subset can be identified with a cover by dyadic subarrays. Looking down through the stack of arrays, this cover gives a tiling of the original $N \times N$ array by square subarrays of size $2^{-m}N \times 2^{-m}N$,where m is the level from which that particular subspace was chosen.

Levels map Here 2 arrays are used. One describes, what is called a "levels map" while the other describes a "coefficient list". Levels map has 2^{2L} integers of length $\log_2(1+L)$ bits each,and describes the level from which a corresponding $N2^{-L} \times N2^{-L}$ block of coefficients in the next list was chosen. The coefficient list consists of $N \times N$ coefficients from the best-basis, scanned in some agreed upon pattern.

If the original signal turns out to be the best-basis representation, then every coefficient will be chosen from level0, the levels map will consists of $2^L \times 2^L$ $0's$, and the coefficients list will contain the original signal. On the other hand, if complete bottom level L turns out to give the best-basis, then levels map will contain $2^L \times 2^L$ $L's$ and the coefficient list will contain all the coefficients from the bottom level, level L. In the case of wavelet basis turning out to be the best, the levels map will contain a description of the wavelet basis and the coefficients list will contain coefficients of the signal.

The estimate overhead cost can be calculated. There are 4^L additional integers each of length $\log_2(1+L)$ bits. This gives a total of $4^L \log_2(1+L)/N^2$ bits per pixel, which can be manipulated by making $L < \log_2 N$. If we start with an image 128×128 and decompose it to level $L = 7$ then the number of overhead bits per pixel is 3.

Subspace lists A basis can also be described by listing the subspaces it contains. One method is to list subspaces by level. 3 arrays are used:

(a) an array num[L] of integers giving the number of subspaces chosen at each level,

(b) an array of arrays subspace[ ][ ], listing the subspace chosen from each level, and

(c) an array containing the complete set of chosen coefficients.

The array num[ ] in (a) contains L entries of varying length, that is, 1 bit for entry 0, which describes whether level 0, $2m$ bits for each entry m which tells as to how

many of the 4^m subspaces on level m are in the best-basis; and so on up to $2L$ bits for level L. Together the subspaces must account for all the coefficients. Hence we have the relation:

$$num[0]*N^2+num[1]*N^2/4+num[2]*N^2/4^2+\ldots+num[L]*N^2/4^L=N^2, \quad (3.1)$$

This implies that $num[L]=4^L(1-num[0]-\ \ldots\ -num[L-1]/4^{L-1})$, so it is not necessary to transmit this value to describe the basis. The total number of bits in the array num[ ] is thus

$$1+2\sum_{m=1}^{L-1} m = 1+(L-1)L = L^2-L+1 \quad (3.2)$$

The array subspace[ ] of arrays in (b) contains L entries of length num[0],num[1] 1], ...,num[L-1], respectively. Array subspace[m][ ] contains integers of length $2m$ bits; array subspace[0] need not be allocated, since there is unique subspace at level 0 Similarly array subspace[L] need not be allocated. Suppose the subspace numbering scheme assigns the indices $4k, 4k+1, 4k+2$, and $4k+3$ to the subspaces descended from $4k$. The actual subspaces in level L be labelled by integers $0,\ldots,4^L$, and the ones that are actually present are the survivors after indices $4^{L-m}k,\ldots,4^{L-m}(k+1)-1$ are deleted for each subspace k at level $0 \le m < L$.

With this scheme the total number of bits in subspace[ ][ ] is

$$\sum_{m=1}^{L-1} 2m*num[m] = 2(L-1)*4^{L-1}$$

Hence the overhead for this coding scheme requires at the most

$$(L^2-L+1+2(L-1)4^{L-1})/N^2 \quad (3.3)$$

bits per pixel. This can be controlled by limiting L. In case of a 128×128 image analysed down to level 7 this value is 3.

3.3 Reconstruction

A picture represented as coefficients may be reconstructed by calculating the value at each point of the appropriate linear combination of wavelet packets. This is done as fllows. First we allocate enough memory for the deepest level of the tree of subspaces that contain retained coefficients, and insert the coefficients into their apropriate locations. Then we reconstruct the parent subspaces by applying the adjoints of the convolution and decimation operators, which produces part of the next deepest level. Into this we add the retained coefficients which belong in that level, at their respective locations, and reconstruct the parent at this level. We continue in this manner until the root has been constucted. Fig (2.13) shows reconstruction from one level.

3.4 Complexity of the Algorithm

Suppose S is an N-element picture. Applying convolution-decimations to generate the tree of coefficient sequence requires $O(N \log N)$ opreations. Calculating information costs has complexity $O(N \log N)$. Labelling "kept" subspaces is equivalent to a breadth-first search through the tree, which has complexity $O(N)$. Locating topmost "kept" subspaces is equivalent to depth-first search, with complexity of $O(N)$, and filling an output register with coefficients from the best-basis takes an additional $O(N)$ operations. We can then perform radix sort to determine the largest coefficients in the utput register, which has a complexity of $O(N \log N)$.

Reconstruction from the retained coefficients has the same complexity as generating all the coefficients, since the requirement of reproducing the entire tree still remains.

3.5 Results of Image Compression

3.5.1 Determination of Compression Factor

The compression factor is determined in the following manner:-

(1) First we calculate the extra number of bits needed to send the additional information by Eqn (3.3).

(2) Then we find the max value of energy and accordingly decide on a threshold.

(3) We then discard the subspaces having energy less than the above threshold.

(4) Now, the total number of bits per pixel, if the image was not to be compressed as above depends on the number of quantization levels Q and is equal to 8 if the quantization levels are 256 ($\log_2 Q$). Thus, the total number of bits to be transmitted in this case is equal to total number of data samples multiplied by number of bits per pixel.

(5) The total number of bits needed for transmission, if the image were to be compressed, using best basis depends on the number of samples retained after steps 2 and 3 above, and the extra number of bits per sample required as per step 1 above. Numberof bits per sample , then, is equal to the number of bits per sample obtained through quantization, (as in step 4) and the extra bits per sample required, for transmitting additional information. This figure is then multiplied by the number of samples retained.

(6) Hence the compression is defined as :-

$$\text{Compression Factor} = \frac{\text{Total No of bits obtained in step 5}}{\text{Total No of bits obtained in step 4}}$$

3.5.2 Results

For the purposes of testing the above results, three images were used, viz, Baboon, Nutan and Lenna.

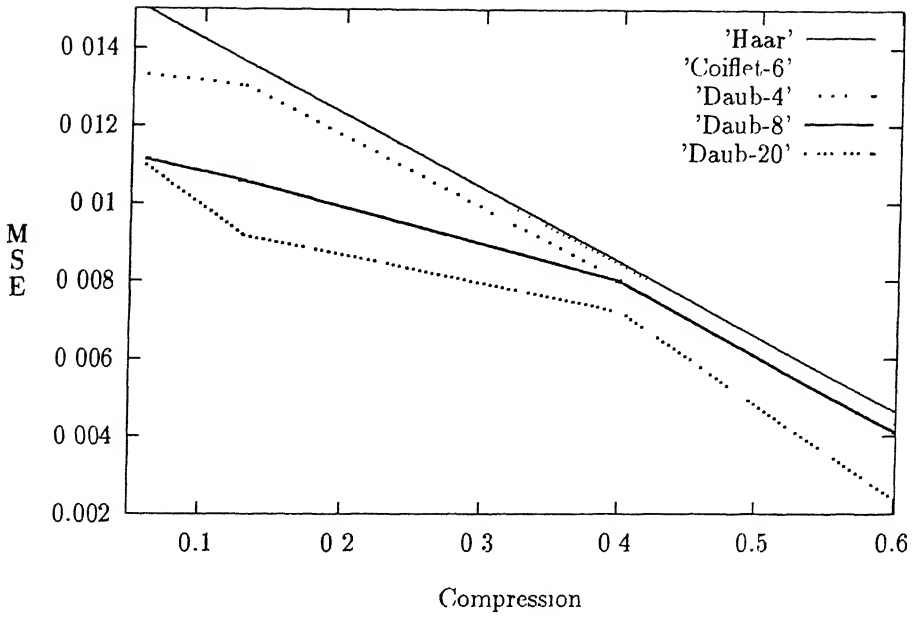


Figure 3.3: Plot of Compression achieved Vs Mean Square Error for various wavelets.

It is seen that, the compression factor varies with the type of image selected.

The results thus achieved have been tabulated in Tables (3 1) ans (3.2). PSNR mentioned above is calculated as under:-

$$\text{PSNR} = -20 \log_{10} \left(\frac{\text{rms error}}{2^n - 1} \right)$$

where n is the number of bits per pixel. Plate (1) shows the original Baboon image. Plates (2), (3) and (4) show reconstructed image using Daub-20, Coiflet-6 and Haar filter respectively.

The results achieved depend on the type of wavelets used. The comparative results of compression achieved by using different types of mother wavelet are shown in Fig. (3.3. The criteria for the selection of these wavelets has been discussed in section (3.10). The wavelet function and scaling function for these wavelets has been given in section (3.10). At this stage we will suffice by tabulating regularity and variance in respect of various types of wavelets used and observing the fact that regularity and variance do have a direct bearing on the image reconstruction and hence on compression achieved. The detailed discussion is postponed till the above quoted reference.

Filter	Compression	Mean Square Error	PSNR
Haar	0.06	0.015	30.5 dB
	0.20	0.01074	31.9 dB
	0.68	0.004645	35.93 dB
Coiflet-6	0.06	0.01422	30.68 dB
	0.12	0.013676	31.01 dB
	0.41	0.008203	33.10 dB
	0.67	0.004305	35.8 dB
Daub-4	0.06	0.0133	30.96 dB
	0.13	0.01303	31.06 dB
	0.40	0.008016	33.17 dB
Daub-8	0.06	0.011142	31.74 dB
	0.13	0.0.010591	31.96 dB
	0.40	0.007979	33.11 dB
	0.64	0.004102	36.03 dB
Daub-20	0.06	0.01098	32.11 dB
	0.13	0.009161	32.75 dB
	0.40	0.007210	33.65 dB
	0.69	0.002360	36.9 dB

Table 3.1: Results for the compression achieved. Image used is Baboon

Filter	Compression	Mean Square Error	PSNR
Haar	0.14	0.00247	36.6 dB
Coiflet-6	0.06	0.008649	31.32 dB
	0.13	0.006338	32.66 dB
	0.45	0.002360	36.96 dB
Daub 4	0.06	0.006251	32.72 dB
	0.13	0.00425	34.22 dB
	0.43	0.001901	37.89 dB
Daub 8	0.06	0.005558	32.85 dB
	0.13	0.005069	34.99 dB
	0.44	0.001790	38.15 dB
Daub-20	0.05	0.007655	31.84 dB
	0.13	0.003793	34.89 dB
	0.41	0.001527	38.84 dB

Table 3.2: Results for the compression achieved. Image used is Nutan

Table (3.3) compares regularity and spatial variance for various types of wavelets.

	Haar	C-6	D-4	D-8	D-20
Regularity	1.0	1.22	1.41	1.7755	2.64
Variance of ϕ	0.0832	0.0986	0.1416	0.179	0.2192

Table 3.3: The comparative table of the various types of wavelets used.

3.6 Edge Charecterisation

Sharp transitions in images are preserved and depicted extremely well in wavelet expansion. Hence edges can be located very effectively as they correspond to sharp transitions. Edge characterisation has two aspects, viz, edge enhancement and edge detection. In wavelet packets representation, as opposed to the wavelet representation, the detail signal is also further analysed along with the average signal. This results in increased flexibility and greater manoeuvreability of the edge information. This feature of wavelet packets decomposition can be exploited for the purpose of edge charecterisation.

Let us assume that a picture S has a resolution L. Let (n, m, k) be the index of an amplitude in the complete wavelet packet expansion, where, $m = 0, 1, \ldots L$. Suppose we wish to detect an edge of scale m_0 in this picture, where a white region darkens to black in a distance $2^{m_0 - L}$. Such an edge will contribute large amplitudes, to scales, $1, 2, \ldots m_0$ at high frequencies. Such an edge can be graphed by selecting only those amplitudes $c(n, m, k)$ above a threshold, sufficiently large, with $m < m_0$ and n greater than an appropriate monotone function of m. In other words, we simply analyse the image down to the desired level L. Then we apply the above threshold to all the subspaces corresponding to the detail signal at that level. And then the signal is reconstructed with the help of the modified subspaces. If only edge

information is required then the average signal can be neglected during the process of reconstruction. Plate (5) shows edge of Lenna image detected in similar fashion.

Edge enhancement can also be achieved in the similar fashion Here we analyse the signal as above. Now we suitably rescale all the subspaces corresponding corresponding to the detail signal while leaving the average signal unchanged. We can then reconstruct the signal using modified subspaces.

Along with the flexibility of selecting the resolution level, (which is also a feature of wavelet basis), wavelet packets also provide the flexibility of selection of subspaces at each level. Thus, depending on the charecterstics of the edges required as also on the number of edges desired, we can select the resolution level, subspace and the threshold for edge charecterisation

3.7 Noise Removal

Denoising is a major problem in image processing. In this thesis, we make use of a visual analysis tool for the purpose of denoising. The visual tool used here is the subspace energy display of the noise and the noisy signal. This method is discused in the next subsection.

3.7.1 Noise Removal using Subspace Energy Display

If the noise that is likely to effect the image can be modeled then it is possible to reduce the noise from the 2-D signal affected by the above noise. The 'noise' can analysed using wavelet packets and the energy distribution in various subspaces at various levels can be displayed. The 'noisy' image that is received can also be analysed in the similar fashion. From the visual information about the noise distribution available from the display above, the subspaces that correspond to more noise energy can be found out. The corresponding subspaces can be either dropped or we can apply a thrshold so as to neglect the coefficients that are likely to contribute towards noise. The algorithm for selection of threshold is discussed in the next subsection. The above energy displays also give a clear picture of the comparative

distribution of energy in various levels. We can then select only that level where subspaces with higher proportion of noise energy correspond to that with lower proportion of signal energy and then carry out the above thresholding. The level so selected is then used for the purpose of reconstruction. As will be seen in the results, an SNR improvement of about 6dB is very easily possible. Since, the above display uses varying grey level intensity to depict the energy distribution, it facilitates in identification of the noisy subspaces and the resolution level to which they belong.

3.7.2 Use of Hypothesis Testing Technique for Calculation of Threshold

The procedure followed above relies on the visual methods for determining the threshold. Hypothesis testing technique can also be used for the purposes of determining the above threshold [12]. Consider an information signal , $s(t)$, corrupted by Additive White Gaussia Noise (AWGN), $w(t)$, with variance σ^2. Then our received signal is $r(t) = s(t) + w(t)$ and the discrete equivalent of the same is $r(k) = s(k) + w(k)$. If $r(k)$ has DWT r_n^l where $l \in L$ and L corresponds to maximum level to which the signal is decomposed, and $n \in N$ where N is theset of points available for each level, then, because DWT is linear operation we have

$$r_n^l = s_n^l + w_n^l$$

where s_n^l is DWT of $s(k)$ and w_n^l is DWT of $w(k)$. Then by using Neyman-Pearson criterion we determine a threshold based on the desired p robability of false alarm. The result of hypothesis testing is the detected parameter set r_n^l, where

$$\overline{r_n^l} = \begin{cases} r_n^l & \text{if } \mid r_n^l \mid \geq th(l) \\ 0 & \text{if } \mid r_n^l \mid < th(l) \end{cases}$$

and $th(l)$ is the threshold at level l. As has been shown in [12], low pass filter $h(n)$ increases the mean of the noise at the output while, high pass filter $g(n)$ yields a zero mean. However, the variance of noise through each filter remains same, provided,

$$\sum_n (h(n))^2 = \sum_n (g(n))^2 = 1$$

The above holds true for orthogonal wavelets. In case of non-orthogonal decomposition the variance at each level can be calculated. Let PR_l, PW_l, and PS_l denote received signal energy, noise energy and information signal energy respectively. We may now calculate these values as under:-

$$PR_l = \sum_n (r_n^l)^2 \; ; \; PW_l = \sum_n (w_n^l)^2 \; ; \; PS_l = \sum_n (s_n^l)^2 \tag{3.4}$$

We can now estimate the expected signal energy as

$$E\{P\hat{S}_l\} = PR_l - E\{PW_l\} \tag{3.5}$$

This signal energy estimate is used to determine the hypothesis testing threshold.Different thresholds will detect diferent sets of parameters. We will find a threshold that yields a parameter set with an energy that is equivalent to that of the signal estimate. That is, let

$$P\overline{R_l} = \sum_n (\overline{r_n^l})^2$$

and vary the threshold $th(l)$ until $P\overline{R_l} = E\{P\hat{S}_l\}$. Once this has been calculated we can obtain the signal estimate as

$$\hat{s}(k) = IDWT\{\overline{r_n^l}\} \tag{3.6}$$

The same algorithm can be applied to non orthogonal wavelets, as also to non gaussian noise.

As we have seen earlier also, wavelet packet analysis gives us greater flexibility even for the noise removal. We have more number of subspaces available for manipulation.

3.7.3 Results of Noise Removal

We have used both, 1-D as well as 2-D signals for the purpose of denoising. The results are tabulated in Table (3.4). In the case of 1-D signals (Sine wave and ECG)

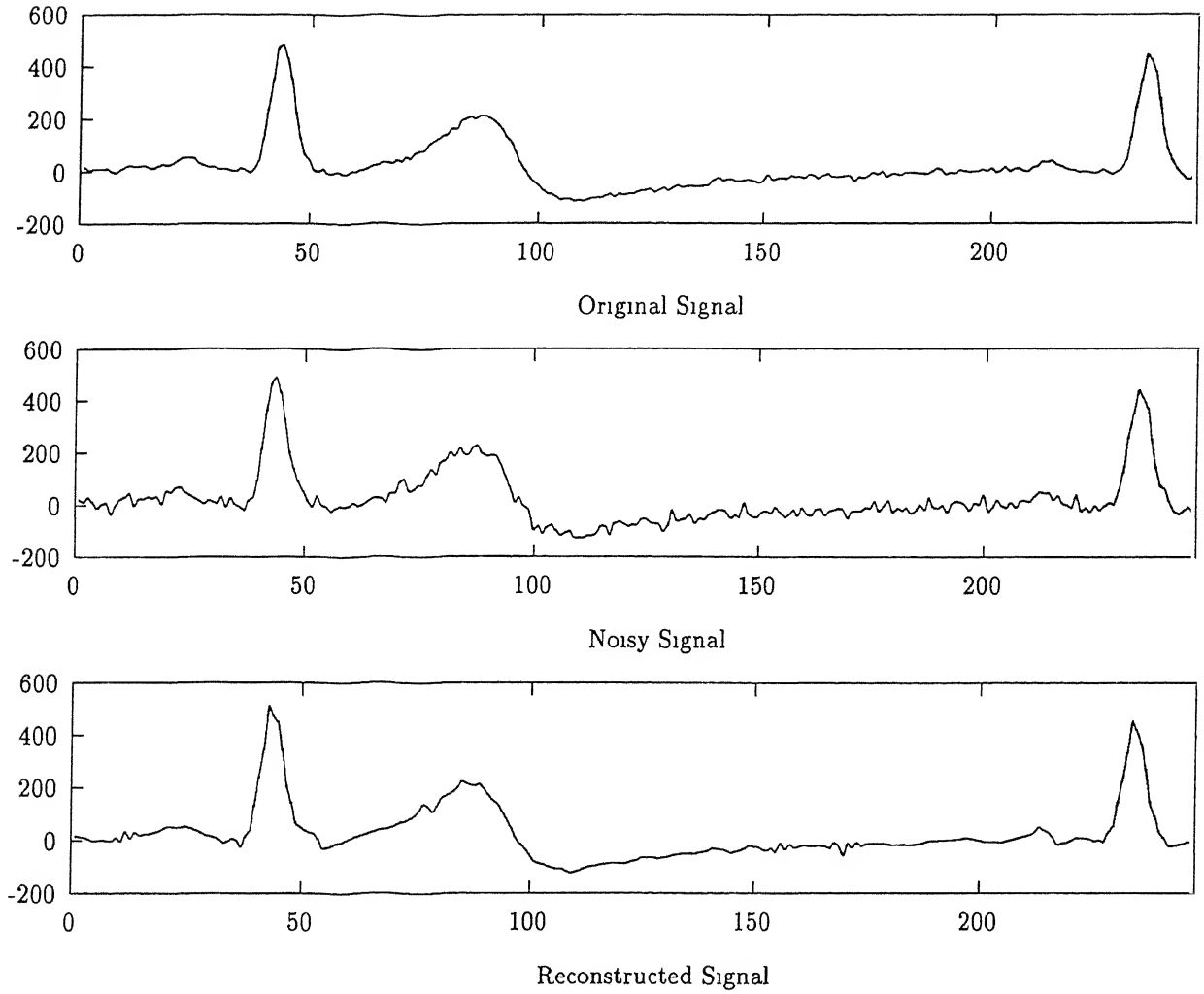


Figure 3.4: Effect of denoising on ECG signal

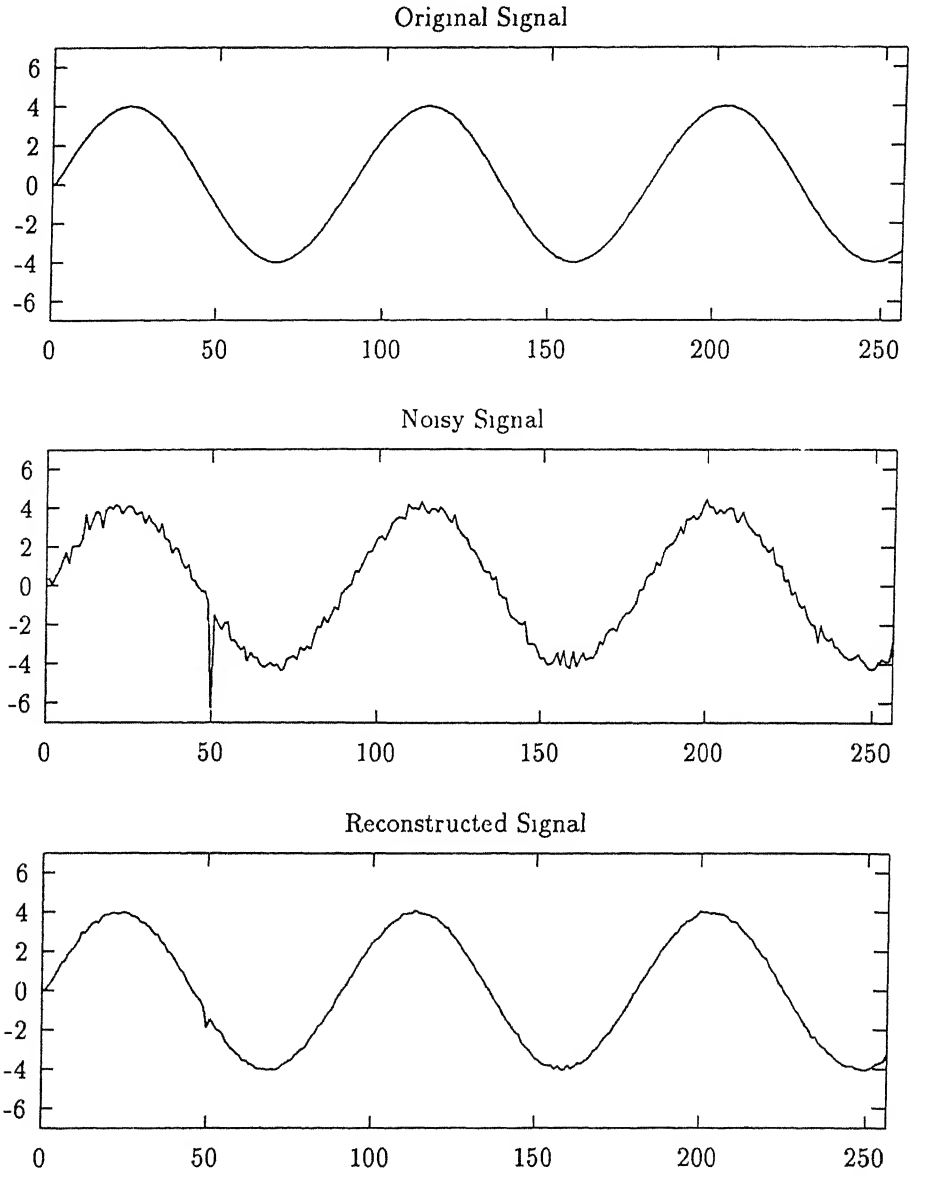


Figure 3.5: Effect of denoising on Sine Wave signal

Signal	SNR before denoising	SNR after denoising	Improvement
Baboon	16.22dB	22.10dB	5.88dB
Lenna	17.65dB	23.66dB	6.01dB
ECG signal	20.12dB	26.67dB	6.55dB
Sine Wave	20.7dB	27.31dB	6.61dB

Table 3.4: Results for Noise Removal.

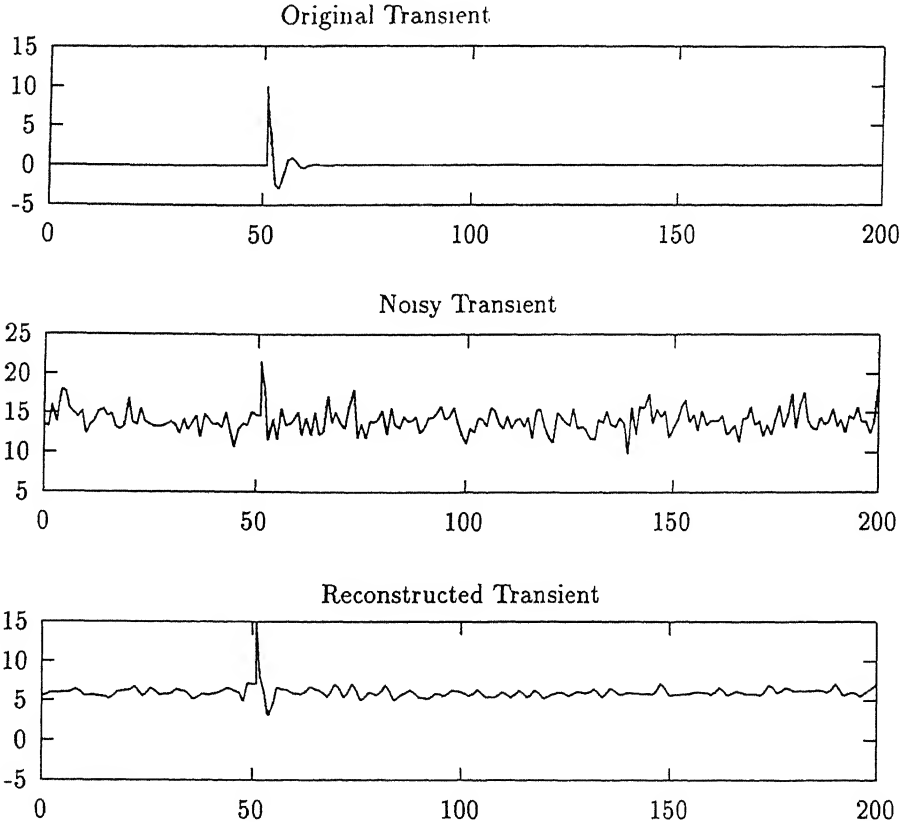


Figure 3.6: Effect of denoising on transient

the same has been shown in Fig (3.4), and Fig (3.5). Plates (6) and (7) show the energy distribution for white noise and 'noisy' Baboon image respectively.i Plate (8) shows 'noisy' Baboon image with a SNR of 16.22dB. Plate (9) shows the effect of denoising. The SNR of this reconstructed image is 22.10dB.

3.8 Local Image Rescaling

For the purpose of highlighting a certain region in the image we resort to local image rescaling. For this purpose, we simply replace $c(n, m, k)$ (notation used is same as that used for edge detection), with $c(n', m', k')$ for a restricted range of $k's$.

3.9 Quality Criteria for Wavelets used in Image Processing

In this section we will try to study certain criteria which would dictate the selection of *wavelets* for processing image signals.

3.9.1 Wavelet Regularity

The regularity of mother wavelet is a very important aspect and is closely related to the regularity of the signal to be processed. Since most of the images are smooth (but for the occasional edges) it looks reasonable to use regular wavelets. Because scaling function $\phi(t)$ and wavelet function $\psi(t)$ are related, they share the same regularity properties. Therefore, in this discussion, we restrict ourselves to the study of $\phi(t)$, which depends only on low pass filter $H(z)$. Zeros at the Nyquist frequency play an important role in determining the regularity. One zero in $H(z)$ at $z = -1$ is necessary to obtain continuity. To achieve a regularity order of N, $H(z)$ must have at least $N+1$ zeros at $Z = -1$. However, if there are other *zeros* also present, then they contribute towards *killing* the effect of regularity caused by zeros at $z = -1$. This effect can be as much as 90%. Rioul [13] studied this effect and derived an

algorithm to estimate the optimal regularity esimate. The same is given below.

(a) Remove all the zeros of $H(z)$ at $z = -1$. These zeros will contribute to a maximum Holder regularity of K, the actual number of zeros at $z = -1$.

(b) The remaining factor $F(z) = (1 + z^{-1})^{-K} H(z)$ will decrease this regularity. The decrease will be by the amount $1 - \alpha$.

(c) α, mentioned above is calculated by first iterating $F(z)$ as under:

$$F^i(z) = F(z)F(z^2)\dots F(z^{2^{i-1}})$$

Associated with this $F^i(z)$, we have time sequence f^i_n. Then, α_i, is calculated as under:

$$\alpha_i = 1/i \log_2 max_n \sum_k \mid f^i_{n+2^i k} \mid$$

(d) Regularity r is then given by:

$$r = K - 1 + \alpha_i \tag{3.7}$$

In practice, i is iterated 20 times to get the regularity estimate. We have already seen regularity estimates of some of the wavelets in Table (3.3). We have observed that higher regularity of the wavelet results in better compression ratio.

3.9.2 Number of Vanishing Moments

Number of *vanishing moments* N, of the wavelet ψ forms an other important criterion. This is a measure of the oscillatory character of the wavelet. The number of vanishing moments is defined as

$$\int x^n \psi(x) dx = 0 \;\; \forall n = 0, 1, \dots, N-1 \tag{3.8}$$

Since ψ is a wavelet, N is greater than or equal to 1 and the property of the wavelets, viz, $\int_{-\infty}^{+\infty} \psi(x)dx = 0$ is verified. In fact, this property of number of vanishing moments is related to regularity of the wavelet. That is, any r-regular multiresolution analysis of the wavelet will generate wavelet ψ with $r + 1$ vanishing

moments. Also, all derivatives of $\Psi(\omega)$ upto the order $N-1$ are zero at the point $\omega = 0$. A wavelet with N vanishing moments enables the cancellation of all wavelet coefficients of a polynomial signal whose degree is less than N. Thus, if f is a polynomial signal of degree less than N, on support of $\psi_{m,n}$ then

$$c_{m,n}(f) = \langle f, \psi_{m,n} \rangle = 0 \tag{3.9}$$

3.9.3 Spatial Charecterization of the Scaling Function

To determine how $\phi(x)$ evolves with respect to x, we introduce the criterion m_k which allows us to charecterize the scaling function ϕ spatially. To define m_k, we introduce the function $p(x)$ such that:

$$p(x) = \frac{\mid \phi(x) \mid^2}{\int \mid \phi(x) \mid^2 dx} \text{ with } p(x) \geq 0 \text{ and } \int p(x)d(x) = 1 \tag{3.10}$$

m_k is then defined as

$$m_k = \int (x - m_1)^k p(x)dx \text{ with } m_1 = \int xp(x)dx \tag{3.11}$$

Thus m_1 is equal to the mathematical expectation of $p(x)$ and m_2 corresponds to the spatial variance of the scaling function ϕ. m_2 allows us to determine the energy concentration of ϕ and provides information on the spatial length or localization of ϕ. This criterion also applys to the wavelet ψ. As we have seen in Table (3.3), along with Tables (3.1) and (3.2) higher spatial variance results in better results. This is because of better energy cocentration associated with the wavelets with higher spatial variance.

Plate 8 :Noisy Baboon Image. SNR is 16.22 dB

Plate 9 :Baboon Image after denoising. SNR is 22.20 dB

Chapter 4

Conclusion and Suggestions for Future Work

4.1 Conclusion

Wavelet Packets have come to be recognised as an important tool in image processing. In this thesis the same have been successfully exploited for feature extraction. Preliminary results obtained are satisfactory, therefore following conclusions can be drawn :

- An image can be analysed using wavelet packets to generate a 'quad-tree' of wavepacket coefficients. From this tree we can select a 'best-basis' which minimizes a certain cost function, e.g. entropy. This representation can then be used to obtain reasonable amount of compression by applying a certain threshold to neglect subspaces whic have insignificant energy.

- The compression achieved in this process depends on the following factors:

  (a) Desired signal-to-noise ratio.

  (b) Type of mother wavelet used. That is, the *regularity* of the wavelet used for the purpose of analysis. It has been observed that higher regularity results in better compression.

- The 'noise' present on the channel can be modeled and analysed using wavelet packets. The signal received on the channel ('noisy signal') can also be analysed in the similar way. The 'quad-tree' obtained above can then be used for generating subspace-energy displays for both, noise as well as the received signal. These displays then form an important visual tool for the purpose of denoising.

- Wavelet packets, provide a wider choice of subspaces. This feature can be effectively used for edge-chareterisation and rescaling a local region in image. By selecting appropriate 'average signal' or 'detail signal', at appropriate level and then rescaling/thresholding we can achieve these aims. The number of edges and their charecteristics depend on the level of image decomposition and the threshold.

4.2 Suggestions for Future Work

It is seen that in image compression, the type of mother wavelet selected has an important effect on the compression ratio achieved. We have used only a few of them. It is suggested that the same algorithm can be tried using some other types of orthogonal wavelets, as well as non-orthogonal wavelets.

In this thesis, while using the concept of multiresolution analysis for the purposes of generation of library of bases, we have always used a resolution factor of 2 (dyadic resolution) along with separable filters. Now the studies have been made using a resolution factor of $\sqrt{2}$ (quincunx resolution). This type of resolution uses non-orthogonal, non-separable filters and will give a finer resolution. Hence, the algorithm discussed in this thesis can be tried using the same.

The package developed for noise removal uses varying gray level intensities to display 'subspace-energy' distribution. The same can be modified to display the energy distribution using different coloures as human eyes are more sensitive to colours. The same package can be made more interactive.

Bibliography

[1] Stephen G. Mallat, "A Theory for zmultiresolution Signal Decomposition; The Wavelet Representation" *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol II, No. 7, July 1989.

[2] Stephen G. Mallat, "Multifrquency Channel Decomposition of Images and Wavelet Models", *IEEE Trans.on ASSP,* Vol.37, No. 12, December 1989.

[3] I. Daubechies, "Orthonormal Basis of Compactly Supported Wavelets" *Comm. in Pure and Applied Math,* Vol. 31, No. 1988.

[4] Mladen Victor Wickerhauser, "Picture Compression by Best Basis Wavelet Packet Coding", Preprint, Yale University, February 1992.

[5] Mladen Victor Wickerhauser, "Accoustic Signal Compression With Wavelet Packets", preprint, Yale University, August 1989.

[6] Ronald R. Coifman, Yves Meyer and Mladen Victor Wickerhauser, "Wavelet Analysis and Signal Processing", Yale University, Connecticut.

[7] , " The Wavelet Transform, Time Frequency Localization and Signal Analysis", *IEEE Trans. on Information Theory,* Vol. 36. No. 5, September 1990.

[8] Ronald R. Coifman and Mladen Victor Wickerhauser, "Entropy Based Algorithm for Best Basis Selection",*IEEE Trans. on Info. Theory, Part-II,* Vol. 38, No. 2, March 1992.

[9] I. Daubechies, *Communications on Pure and Applied Mathematics,* Vol. XIV , No 5, June 1992.

[10] Antonin M. Barlaud, Mathieu P. and I. Daubechies, " Image Coding using Wavelet Transform ", *IEEE Trans. on Image Processing*, Sep 1992.

[11] Yansun Xu, John B. Weaver. Dennis M. Healy and Jian Lu, " Wavelet Transform Domain Filters : A Spatially Selective Noise Filtration Technique",*IEEE Transactions on Image Processing* .Vol 3 No. 6 November 1994.

[12] Joseph P. Noonan, Henry M. Polchlopek, and Michael Vartersian, " A Hypothesis Testing Technique for Wavelet Transform in the Presence of Noise", *Digital Signal Processing* Vol 3, page 89-96. 1993.

[13] Olivier Rioul, "Regular Wavelets: A Discrete-Time Approach",*IEEE Transactions on Signal Processing*, Vol. 41 No. 12 December 1993.